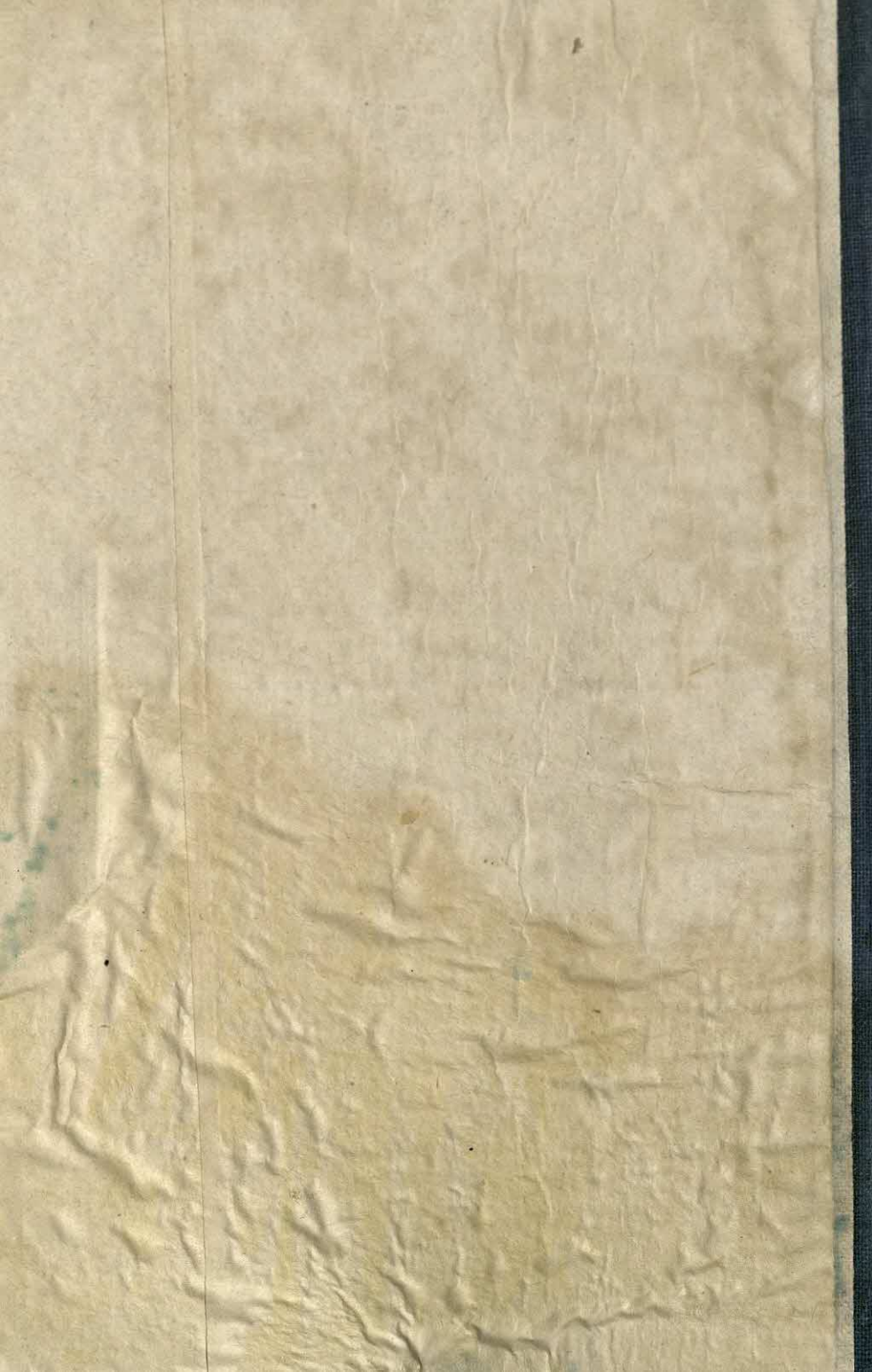






26/9/52
checked
188 m.
Librarian
1/11/72
165

Vol. 85, No. 1

Founded in 1887 by G. Stanley Hall

March 1972

THE AMERICAN JOURNAL OF PSYCHOLOGY

Bureau of Biol. & Psyl. Research

S. C. E. R. T. J.

Date

3.3.81

Acc.

5653

Edited by

Lloyd G. Humphreys
William E. Kappauf

M. E. Bitterman

Benton J. Underwood

-
- JACOB BECK. Similarity Grouping and Peripheral Discriminability under Uncertainty 1
- WILLIAM EPSTEIN. Retention of Sentences, Anomalous Sequences, and Random Sequences 21
- STANLEY C. COLLYER, JOHN JONIDES, and WILLIAM BEVAN. Images as Memory Aids: Is Bizarreness Helpful? 31
- CALVIN D. SOWDER. Interference in Sentence Recall 39
- RICHARD A. KASSCHAU. Polarization in Serial and Paired-Associate Learning 43
- LAWRENCE E. COLE and N. JACK KANAK. Overt/Covert Pronunciation and Rehearsal in Verbal-Discrimination Learning 57

continued on back cover

Vol. 85

1972

Editors Lloyd G. Humphreys M. E. Bitterman
William E. Kappauf Benton J. Underwood

Book Review Editor Julian Hochberg

Consulting Editors William Epstein Robert Fox Wendell Garner
Lawrence E. Jones John Krauskopf Allan Wagner

Index Editor Rand B. Evans Necrology Editor E. R. Hilgard

The *American Journal of Psychology* was founded in the interest of general experimental psychology and publishes reports of original research, shorter notes and discussions, and pertinent book reviews. Manuscripts of articles and notes may be submitted to Drs. Humphreys or Kappauf (University of Illinois at Urbana-Champaign, Champaign, Illinois 61820), Bitterman (University of Hawaii, Honolulu, Hawaii 96822), or Underwood (Northwestern University, Evanston, Illinois 60201). Books intended for review should be sent to Dr. Hochberg (Columbia University, New York, New York 10027).

Authors of articles may order offprints at the time they return galley proof. Other contributors are sent tear-sheet copies as available unless extras are ordered when manuscript is accepted.

Contributors are asked to follow the style described in the *Publication Manual* of the American Psychological Association. Keep abstract and title quite brief, use footnotes only when essential, avoid all but the most common abbreviations (2 mm, 4 deg, 60 db, 15 sec; *but* subjects, experimenter, paired-associate learning), and use the tables and figures to supplement rather than repeat the text or each other; a well-designed table or figure can provide more information than many words of text. If you elect to omit from your manuscript any information that is critical to the evaluation of your data (analyses of variance summary tables, means, variances, correlation and factor matrices), provide two copies of that material for the editors' and their consultants' use in reviewing the paper.

The subscription price is \$10.00 a year (foreign \$10.50); single copies cost \$2.75. Back numbers and volumes may be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which all business inquiries (subscriptions, advertisements, changes of address) should be addressed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

26 SEP 1972

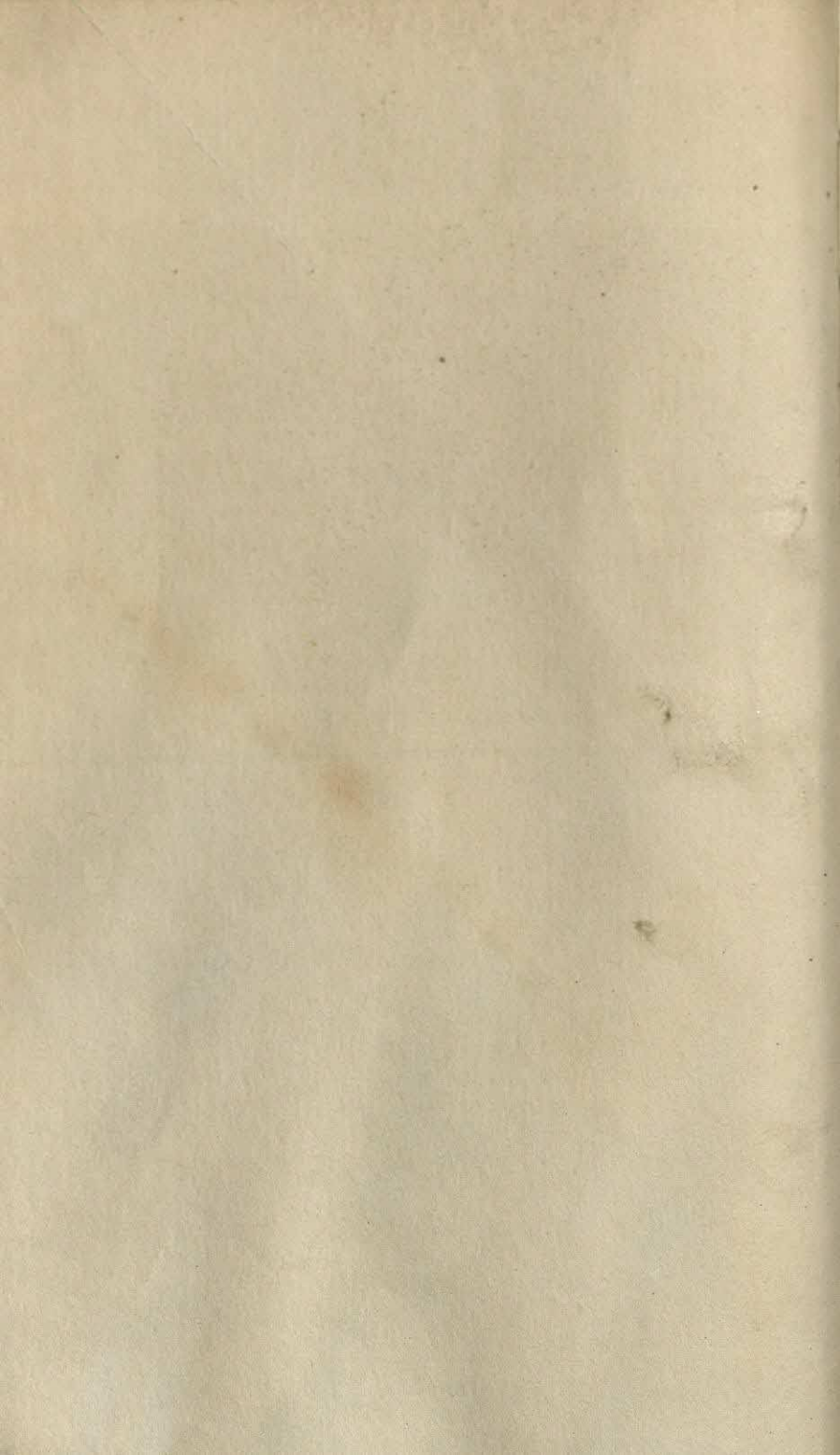
Vol. 85, No. 1

continued from front cover

March 1972

JOHN L. LOCKE and FRED S. FEHR. Subvocalization of Heard or Seen Words prior to Spoken or Written Recall	63
JOHN H. MUELLER and JOHN L. FLANAGAN. Total Time in Verbal-Discrimination Learning	69
RICHARD B. MILLWARD and ARTHUR S. REBER. Probability Learning: Contingent-Event Schedules with Lags	81
WILLIAM F. ORR, CHARLES F. FEDERSPIEL, and O. DALE MAXWELL. Patterns in Random Choice	99
DOUGLAS L. MEDIN. Evidence for Short- and Long-term Memory in Monkeys	117
DONALD G. MACKEY. Input Testing in the Detection of Misspellings	121
ISAAC BEHAR. The Dancing Arabesque: An Unusual Visual Effect	129
Frederic Charles Bartlett: 1886-1969 (R. C. Oldfield)	133
Book Reviews (Lloyd Kaufman, Vincent M. LoLordo, Gerrit Wolf; books received)	141

PUBLISHED QUARTERLY BY THE UNIVERSITY OF ILLINOIS PRESS
Urbana, Illinois 61801



SIMILARITY GROUPING AND PERIPHERAL DISCRIMINABILITY UNDER UNCERTAINTY

Jacob Beck

University of Oregon

The effectiveness of rotated figures in producing similarity grouping corresponded to their peripheral discriminability under uncertainty in a patterned visual field—but not to their discriminability in an empty visual field. The results support the hypothesis that similarity grouping is a segregative process based on the response of the visual system to stimulus differences before attention narrows or focuses.

Beck (1966a, 1966b, 1967) has shown that differences in the *slopes* of the component lines of figures are a major variable in similarity grouping. Changes in the orientation or shape of some of the figures in a field, changes that altered the slopes of their component lines from vertical (90 deg) and horizontal (0 deg) to 45 and 135 deg, facilitated segregating the field into distinct perceptual groups more readily than changes in the orientation or shape of these figures that left the slopes of their component lines vertical and horizontal.

The *arrangement* or configuration of the component lines of figures may also affect similarity grouping. Disparate L figures in a pattern of upright Ts grouped better than other figures with vertical and horizontal lines, such as a T on a side or inverted, or a + (Beck, 1967). Similarly, not all figures in which the slopes of their component lines were altered to 45 and 135 deg grouped equally well. Disparate X figures in a pattern of upright Ts did not group nearly as well as did disparate Ts tilted 45 deg, even though the slopes of the component lines of these two disparate figures were the same (Beck, 1966a; 1967).

In an attempt to characterize the processes underlying visual organization, Beck (1967) proposed that it is the figural properties selectively responded to before focal attention that produce strong grouping. The term 'focal attention' is vague and has been applied to a number of different processes that are often not distinguished; here, however, it refers to the degree or concentration of attention. Operationally, the hypothesis is that the figural properties producing similarity grouping correspond to those that are readily discriminated when figures are presented periph-

erally in a patterned visual field in which an individual does not know where to direct his attention—correspond, that is, to the properties that produce peripheral discriminability under uncertainty. The experiments to be reported test that hypothesis.

Two studies were undertaken. Study I presents the results of three experiments assessing the degree of grouping produced by six different pairs of figures. Study II reports the results of experiments on the peripheral discriminability of the same figures in a patterned visual field and in an empty visual field. Because of the presence of other figures in a patterned visual field, the brightness difference between figure and background cannot direct a subject's attention to the figure to be discriminated—the uncertainty situation. Conversely, in an otherwise empty visual field, the brightness difference between a figure and its background *can* direct a subject where to attend, and discriminability thus becomes a function also of the processes of analysis operating when attention is focused. According to the hypothesis proposed, the effectiveness of figural properties in producing similarity grouping in Study I should correspond to their discriminability in the patterned visual field, but not the empty visual field, of Study II.

STUDY I

Three experiments assessed grouping. In Experiment I the time taken to count disparate U and V figures embedded in larger matrices was used as an indirect index of perceptual grouping. The rationale for this measure was that grouping makes the disparate figures stand out and should decrease the time required to count them. In Experiments II and III the grouping of these same figures was assessed directly by means of ratings. Eighteen subjects served in Experiments I and II, and 15 served in Experiment III, different subjects in each experiment. They were all paid volunteers and were naive as to the purpose of the experiments.

Method

—Experiment I—The stimuli consisted of 2¼-by-3-in. displays of background and disparate figures. Six combinations of background and disparate figures were used (Figure 1). The rotations from background to disparate figures were 45 deg counterclockwise ('tilted') for display A, 90 deg counterclockwise ('sideways') for display B, 180 deg ('inverted') for display C, 45 deg clockwise ('tilted,' again) for display D, 90 deg counterclockwise ('sideways,' again) for display E, and 180 deg ('inverted,' again) for display F. The background and disparate figures in displays A and D differ both in the slopes of their com-

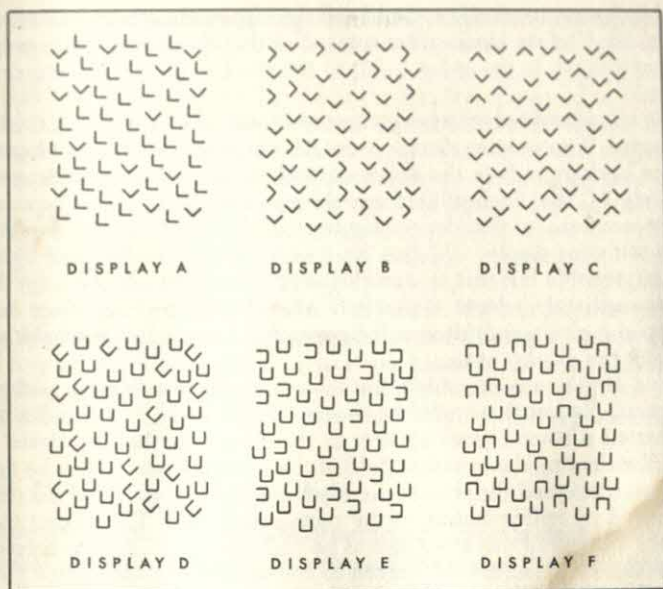


Fig. 1. Sample displays used in Experiments I and II and III

ponent lines and in their orientations, and those in displays B and E and C and F have the same slopes of their component lines but differ in their orientations. Note, however, that inverting the disparate figures leaves them symmetric about the vertical median in displays C and F, whereas turning them on a side changes their axes of symmetry from a vertical to a horizontal median in displays B and E.

Each of the six combinations of figures was drawn in india ink in six different matrix arrangements. In each matrix there were 56 figures arranged in seven columns and eight rows. The figures were drawn misaligned with each other in order to avoid introducing colinear lines. The disparate figures were placed randomly throughout the matrix; there were never more than three in one row or four in one column. Two matrices consisted of 42 background figures and 14 disparate figures, two of 41 background and 15 disparate figures, and two of 40 background and 16 disparate figures. The two lines composing each figure were $\frac{3}{16}$ in. long and were always at right angles to each other. The lines of the figures were always either vertical and horizontal or tilted at 45 and 135 deg.

The six combinations of background and disparate figures drawn in the six matrix arrangements made a total of 36 stimulus displays. These were photographed and printed on high-contrast paper. The prints were centered and mounted on individual pieces of white cardboard. The 36 displays were then arranged in blocks of six; each combination of background and disparate figure and each matrix arrangement appeared once in each block. The displays

in each block were randomly ordered and the displays numbered consecutively from 1 to 36. The six blocks were rotated so that the stimuli were presented to the first subject in the order 1-36, to the second subject in the order 7-6, and so on.

The 36 stimulus displays were presented by means of a two-field Gerbrands tachistoscope. The viewing distance was 23.5 in., and the mean luminances of the white backgrounds in the exposure and the preexposure fields were approximately 3.5 ftL. Each subject was shown examples of the six types of displays. He was told (a) that depressing the switch in front of him would make visible a stimulus display, (b) that he was to count the disparate figures as quickly as possible but that it was also important to be accurate, (c) that he was to immediately release the switch when he finished counting, because this stopped a clock, and (d) that if he counted incorrectly he would be presented with the display again.

Before a display was presented, the subject was always told the background and disparate figures that would be shown. When ready, he pressed a switch, which started a Hunter clock counter at the same time that it exposed a display by illuminating the exposure field of the tachistoscope. When he finished counting and released the switch, the clock counter stopped, and the stimulus was removed by reillumination of the preexposure field. The subject then reported the number of disparate figures he had counted. If he was correct, the experimenter recorded the time taken to count the disparate figures (in hundredths of a second). If incorrect, the subject was told that he had not counted correctly, and the stimulus was presented again at a later point in the series. In order to become familiar with the experimental task, each subject received two preliminary trials with practice displays not used in the experiment.

—Experiment II—The same six combinations of background and disparate figures were scaled for grouping. A Gerbrands tachistoscope was again used to present the stimulus displays. Twenty-four of the 36 displays used in Experiment I were presented in a random order, four of each of the six types of displays. The random order was constrained so that no more than two displays with the same combination of background and disparate figures appeared in a block of eight trials. The exposure duration was 4 sec. The subject was asked to scale how well the disparate figures stood out immediately, as a whole, without his having to search the display for them. He was instructed to indicate the degree to which the background and disparate figures were seen as separate groups on a seven-point rating scale, where 0 indicated that the disparate figures did not stand out clearly from the background figures (that the background and disparate figures seemed to compose one pattern) and 6 indicated that the disparate and background figures were clearly seen as two distinct groups. Care was taken to ensure that each subject understood what was to be judged. He was given examples of displays that grouped well and weakly on the basis of differences of length of lines. In order to familiarize him with the stimuli to be scaled, the entire set of stimuli was shown before the experiment began.

—Experiment III—Four of the matrices prepared for Experiment I were mounted in a square arrangement on white cardboard. The combinations of background and disparate figures were those of displays A, B, C, and E (see

Figure 1). The four matrices in a square arrangement were photographed and printed on high-contrast paper in ten sizes. The sizes of the individual displays ranged from $\frac{3}{8}$ by $\frac{3}{4}$ in. to $2\frac{3}{4}$ by 3 in. in ten equal steps. The lengths of the two lines making up the figures in the displays increased in $\frac{1}{64}$ -in. steps from $\frac{3}{64}$ in. to $1\frac{3}{64}$ in. A subject viewed the stimulus card at eye level at a distance of 65 in. He was instructed to rate, on a scale from 0 to 100, how well the disparate figures in each of the four displays on a stimulus card grouped and were seen as separate from the background figures. The instructions and examples of grouping given were the same as in Experiment II, except that the 0-to-100 scale was used. After each stimulus presentation, the subject turned around and inspected a lithograph print, during which time a new stimulus card was inserted in the stimulus holder. The ten stimulus cards containing the four combinations of background and disparate figures were presented four times each for a total of 40 presentations. The sequence of stimuli was randomized differently for each subject. To familiarize a subject with the stimuli to be scaled, the entire set was shown before the experiment began.

Results and discussion

Figure 2 summarizes the data obtained in Experiments I and II. The first figure in each pair of figures shown illustrates the background figure, and the second the disparate figure. The number below each pair of figures in Experiment I is the mean time in seconds taken to count the disparate figures. The number below each pair of figures in Experiment II is the mean rating of the degree to which the background and disparate figures were perceived to group. The means shown were averaged over subjects and stimulus displays for each type of combination of background and disparate figures. The stimuli are arranged from the greatest to the least amount of grouping.

In Experiment I grouping was assumed to be inversely related to the time taken to count the disparate figures. The correlation between those mean counting times in Experiment I and the mean ratings for grouping in Experiment II was in fact $r = -.95$. The relationship was remarkably stable over trial blocks. Although the counting times tended to decrease from the first trial block to the sixth, the correlations between mean counting times and mean ratings for grouping calculated for each of the six trial blocks ranged between $r = -.94$ and $r = -.96$.

In order to compare the degree to which the different combinations of background and disparate figures produced grouping, a Tukey test of multiple comparisons was applied to the mean counting times and the mean ratings. Inspection of the mean counting times revealed that the sample means were approximately proportional to the standard deviations for the respective combinations of background and disparate figures. The Tukey test was, therefore, applied to the logarithms of the

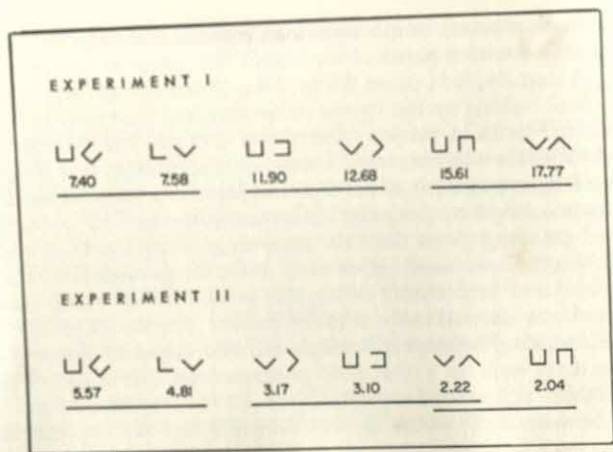


Fig. 2. The mean time to count the disparate figures in Experiment I and the mean ratings of grouping in Experiment II; the first figure in each pair indicates the background figures in a display, the second one indicates the disparate figures

counting times. The mean counting times and the mean ratings that were not significantly different from each other ($p < .01$) are underscored by the same lines in Figure 2.¹ In both the counting and scaling experiments, the strongest grouping occurred with the tilted disparate figures in displays A and D, intermediate degrees of grouping with the sideways disparate figures in displays B and E, and the least amount of grouping with the inverted disparate figures in displays C and F.

Thus, as in previous studies (Beck, 1966a, 1966b, 1967; Olson and Attneave, 1970), changes in orientation that altered the slopes of the component lines of the disparate figures gave the best grouping. On the other hand, the sideways disparate figures in displays B and E produced better grouping than the inverted disparate figures in displays C and F, though those sideways and inverted figures all had the same slopes of their component lines as the background figures. As will be indicated below, this too may be due to differences in the slope properties of the sideways figures, rather than to the difference in their axes of symmetry.

The mean ratings of grouping as a function of figure size for the four combinations of background and disparate figures in Experiment III are shown in Figure 3. An analysis of variance revealed that the factors of figure size and figure combination were statistically significant [$F(9, 126) = 52.42, p < .001$; $F(3, 42) = 17.03, p < .001$], as was their interaction

[$F(27, 378) = 6.61, p < .001$]. The mean ratings obtained when the component lines of the figures were $\frac{1}{4}$ in. long are comparable to the results obtained in Experiments I and II. The best grouping occurred with the tilted figures in display A, intermediate and approximately equal degrees of grouping occurred with the sideways figures in displays B and E, and the least grouping occurred with the inverted figures in display C. Statistical tests of the differences between displays A and B [$t(14) = 3.10$], A and E [$t(14) = 6.22$], B and C [$t(14) = 3.81$], and E and C [$t(14) = 3.47$] were significant ($p < .01$).

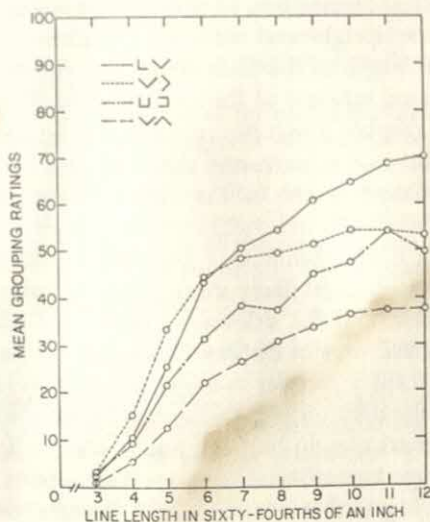


Fig. 3. The mean ratings of grouping as a function of figure size in Experiment III; the first figure in each pair of figures indicates the background figures in a display, the second one shows the disparate figures

Since grouping occurred weakly or not at all for all four displays when the lines composing the figures were less than $\frac{1}{4}$ in. long, a significant interaction between figure size and figure combination is to be expected. Of particular interest, however, is the interaction between figure size and figure combination for displays A and B and for displays B and E. Inspection of Figure 3 indicates that for line lengths greater than $\frac{1}{4}$ in., the tilted figures in display A tended to group better than the sideways figures in display B, while for line lengths less than $\frac{1}{4}$ in., the sideways figures in display B grouped better. The sideways Vs in display B also

grouped markedly better than the sideways Us in display E for line lengths between $\frac{5}{64}$ and $\frac{8}{64}$ in. Analyses of variance revealed that the interaction of figure size and figure combination was statistically significant for displays A and B [$F(9, 126) = 7.35, p < .001$] and for displays B and E [$F(9, 126) = 2.28, p < .025$]. The results indicate that the relative grouping of display B improved with smaller line lengths.

Phenomenally, the salience of the overall orientations of the sideways V figures in display B appears to have been greater when line lengths were small than when large. The upright figures when small appeared to have an overall horizontal orientation, and the ones turned on a side, to have an overall vertical orientation. In contrast, the salience of the overall orientations of the upright and sideways U figures in display E did not change with the length of the lines composing the figures. It is possible that the increased salience of the overall orientations of the small disparate figures in display B was the result of this upright figure's having its greatest extent in the horizontal direction and this sideways figure's having its greatest extent in the vertical direction. Olson and Attneave (1970) have pointed out that in terms of slope analyzers, an upright V would most likely stimulate a horizontally oriented receptive field and on its side would most likely stimulate a vertically oriented field. It is therefore possible that the sideways figures in display B grouped better than the inverted ones in display C because the background and disparate figures in display B differ in the slope analyzers they stimulate and not because of the difference in their axes of symmetry.

Figural properties are closely coupled, and rotating a figure will often simultaneously change a number of form variables. For example, rotating an upright U 90 deg, so that it is on its side, changes both the orientation of the figure and the number of horizontal and vertical lines. The better grouping of the sideways figures in display E in comparison with the inverted ones in display F may also be because the background and disparate figures in display E differ in the number of their vertical and horizontal lines rather than because of the difference in their axes of symmetry.² Because of this close coupling of figural properties, similarity grouping, which appears to be due to relational form variables such as line arrangement or figural orientation, may actually be the result of differences in the slope and brightness properties of the figures.

STUDY II

Study II was concerned with peripheral discriminability under uncertainty. The experiments to be reported were done with the U and V fig-

ures used in Study I. A large number of other experiments, however, have been conducted on peripheral discriminability and their results will be introduced into the discussion where appropriate.

Method

—Experiments IV, V, VI, and VII—The patterned visual field consisted of four figures in the corners of an imaginary square. Each target consisted of either all upright figures or three upright figures and a disparate figure. The disparate figure was a rotation of the upright figure and could appear in any of the four corners; it could be rotated 45 deg clockwise to form a 'tilted' figure, rotated 90 deg clockwise to form a 'sideways' figure, or rotated 180 deg to form an 'inverted' figure. The center-to-center distances along the sides of the imaginary square were 3 in. in Experiments IV and VI, and 2 in. in Experiments V and VII.

Two sets of targets were prepared for each experiment. Each set consisted of ten targets. One set had two targets in which all of the figures were upright, four targets in which the disparate figure was an inverted one, and four targets in which the disparate figure was a tilted one; in the other set, a sideways figure was substituted for the tilted one. The exact shapes of the upright and disparate figures composing each of the eight sets are shown in the column heads of Figure 4. The order of presenting the two sets in an experiment was alternated from subject to subject. A 10-min interval separated the presentations of the two sets of targets.

The figures were drawn in india ink on heavy white paper. The lines composing each figure were $\frac{3}{16}$ in. long and were always at right angles to each other. The targets were mounted on white cardboard and presented by means of a Gerbrands tachistoscope. The backgrounds of the preexposure and exposure fields were white with average luminances of 3.5 ftL. A black dot placed in the center of the preexposure field served as a fixation point. A dot was also placed in the center of the square array on each target. Each target was carefully aligned so that when the target was flashed, the fixation point in the preexposure field coincided exactly with the dot in the center of the square array of figures on the target. The distances between the center dot and the figures in the corners were 2.1 in. for the figures arranged in a 3-in. square (5 deg 7 min of arc) and 1.4 in. for the figures arranged in a 2-in. square (3 deg 26 min).

Separate groups of 20 subjects served in each of the four experiments. The targets were flashed at $\frac{1}{50}$ of a sec. The subject first was shown examples of the targets in a set and then asked to indicate whether a flashed target had all identical figures or, if it had a disparate figure, to identify the figure and its location ('tilted, upper left'). He was told that a disparate figure could appear in any of the four corners and that fixating the center dot would enable him to perform optimally.

Each set of ten targets was presented four times; each of the two types of disparate figures in a set, 16 times. The sequence of stimuli was randomized differently for each subject with the constraint that all ten of the targets appear once in every block of ten trials. Practice trials with targets differing

from those used in his experiment were given to familiarize him with the task. He was not told either the number of times the different types of targets would be presented or the correctness of his responses. The experimenter gave a warning signal before presenting a target. The subjects were naive as to the purposes of an experiment, and all had normal vision or corrected-to-normal vision with glasses. In exploratory experiments, people who wore contact lenses were found to perform poorly and were therefore not run.

—Experiments VIII and IX—In an otherwise empty visual field, a single U figure appeared on one corner of an imaginary 3-in. square. The distances from the fixation point to the target were all 2.1 in. (5 deg 7 min of arc). Sixteen targets were prepared. Four of the targets had an upright figure, four had a 45-deg-tilted figure, four had a sideways figure, and four had an inverted figure. Each figure was composed of lines $\frac{3}{16}$ in. long. The stimuli were presented in three sets. Each set consisted of four targets with a single upright figure appearing once in each of the four corners of the imaginary square, and four targets each with a single tilted, sideways, or inverted figure appearing once in each of the four corners. The tilted, sideways, and inverted figures were always presented in separate sets. The exact shapes of the two figures making up each set are shown in the column heads in Figure 5.

The subject was told before each set the two figures (upright and, say, tilted) that would be presented in that set. Each of the three sets was presented twice in an experiment. There were 11 stimulus presentations in a block of trials. Three of the eight stimuli in a set were randomly added to a block of trials but were not analyzed. Thus, each of the two types of figures in a set occurred in eight analyzed trials. The sets were presented in a different irregular order to each subject. The general procedure and apparatus was the same as in the experiments in a patterned visual field. Each subject was again instructed to fixate a central fixation point and told that this would enable him to perform optimally. To familiarize him with the task, the three sets of targets were presented in a practice session.

The practice session was also used to determine an exposure duration that would give an approximate overall percentage correct. In Experiment VIII the exposure duration for each subject was chosen to give approximately 80% correct responses overall. In Experiment IX the exposure duration was chosen to give each subject approximately 60% correct overall. Fifteen different subjects served in each experiment. They were naive as to the main purposes of the experiment. All had normal vision or corrected-to-normal vision wearing glasses.

Results and discussion

Figure 4 presents the mean number of errors the subjects made with the patterned visual field. An error was counted when a disparate figure was presented and not accurately reported. Inspection of the data revealed that whether a subject was presented first or second with a set of targets had no important effect on the pattern of errors made as a function of the rotation of the disparate figure.

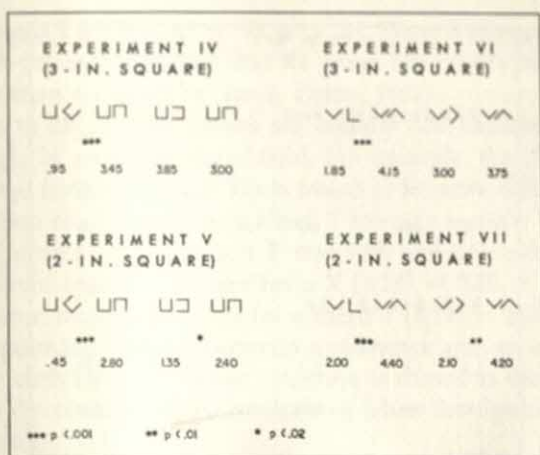


Fig. 4. Mean errors on rotated disparate figures in a patterned visual field in Experiments IV, V, VI, and VII; the first figure in each pair of figures indicates the upright figure on a target, the second one indicates the disparate figure

The results show that changes in figural orientation that altered the slopes of the figure's component lines were more effective in directing the subject's attention to a disparate figure than changes in orientation that left the slopes of the component lines of the figure vertical and horizontal. The results of two-tailed t tests are also shown in Figure 4. With a 3-in. square significantly fewer errors were made when the disparate figure was tilted than when it was inverted. When the disparate figure was sideways, the subjects did not discriminate it significantly better than when it was inverted [$t(19) = 1.37$, $p > .1$; $t(19) = 1.70$; $p > .1$]. With a 2-in. square both the tilted and sideways figures were discriminated significantly better than the inverted figure.

Statistical tests also showed that a tilted U was discriminated significantly better than a sideways one both in a 3-in. square [$t(19) = 4.56$, $p < .001$] and in a 2-in. square [$t(19) = 2.75$, $p < .025$]. A tilted V was discriminated significantly better than a sideways one only in a 3-in. square [$t(19) = 3.38$, $p < .005$]. In general, then, the results indicate that the peripheral discriminability under uncertainty of these two basic figure shapes rotated in a patterned visual field did indeed correspond to their effectiveness in producing similarity grouping.

Figure 5 presents the mean number of errors made when single figures were presented in an empty visual field. The overall percentage correct was 78% in Experiment VIII and 55% in Experiment IX. In

EXPERIMENT VIII					
U	↙	U	⌋	U	⌈
1.87	1.80	1.67	1.73	1.60	1.67
EXPERIMENT IX					
U	↙	U	⌋	U	⌈
3.53	3.40	3.60	3.53	3.60	3.67

Fig. 5. Mean errors on upright and rotated disparate figures in an empty field in Experiments VIII and IX

both Experiments VIII and IX, the mean numbers of errors for a tilted and a sideways figure were similar to each other and the same as for an inverted figure. Figure 5 shows also that the mean number of errors for an upright figure was about the same when it was presented in a set with a tilted figure as with an inverted one. Thus, when a brightness difference between a figure and a background directed a subject where to attend, the relative discriminability of a rotated figure, as hypothesized, did *not* correspond to its effectiveness in producing similarity grouping. The features distinguishing a tilted, sideways, or inverted U from an upright one were equally discriminable when attention was directed to them.

Peripheral discriminability under uncertainty in a patterned visual field—but not in an empty visual field—has also been found to correspond to the grouping of other disparate figures. In a supplemental experiment similar in procedure to Experiments IV–VII, subjects discriminated between targets in which the four T figures in the corners of an imaginary 3-in. square were the same, and targets in which three of the figures were the same (a T) and the fourth (a X or a tilted T) was different. The mean number of errors for X was 6.20 and for tilted T 4.13 [$t(14) = 2.54, p < .025$]. When the subjects then discriminated the same three figures presented singly in an empty visual field, a procedure like that of Experiments VIII and IX, the mean number of errors in discriminating between a X and an upright T was 3.00 and between a tilted

and an upright T 2.87 [$t(14) = .98, p > .2$]. These findings are also consistent with the earlier finding that Xs group less readily in a pattern of upright Ts than do tilted Ts (Beck, 1966a; 1967).

Needless to say, not all figures are equally discriminable when presented singly in an empty visual field. For example, the discrimination of a *sideways* from an upright T was found to be more difficult than the discrimination of either a X or a tilted T from an upright T. The mean number of errors for a sideways T was 5.60 and differed significantly from the mean number of errors for a X [$t(14) = 5.65, p < .001$] and from the mean number of errors for a tilted T [$t(14) = 6.07, p < .001$]. The distinguishing features between a sideways and an upright T are apparently such that even when attention is drawn to the figures, it is difficult to discriminate between the two when the figures are flashed briefly outside the fovea.

Finally, Eichelman, using somewhat different procedures (1970), also found that in an empty visual field there were no significant differences in the discriminability of a tilted T, a X, or a L from an upright T. But adding upright Ts to the background changed the relative discriminability of these figures, a tilted T becoming more discriminable than either a X or a L. Eichelman (1970) also found that when a subject is instructed where to attend in a patterned visual field, the discriminability of a tilted T, a X, and a L were not differentially affected by the introduction of background figures.

GENERAL DISCUSSION

Figural properties

The results of the present experiments are consistent with the working hypothesis that the effectiveness of figural properties (stimulus variables) in producing similarity grouping corresponds to their effectiveness in directing an individual's attention to a peripheral stimulus when he does not know its location. Differences in brightness, movement, color, and the slope of lines are effective in directing an individual's fixations as he searches for a stimulus in a visual field (Hochberg, 1965; Williams, 1967). These variables are also those that are most effective in producing similarity grouping.

The more general hypothesis the studies appear to support is that similarity grouping is fundamentally an acuity task involving the simultaneous discrimination of stimulus differences prior to a narrowing or focusing of attention. According to this theory, similarity grouping is not

the result of a general process of information processing (such as the use of redundancy) to economically encode a visual pattern. Rather, it depends on specific stimulus properties picked up prior to focal attention. For example, differences in direction give strong grouping only when associated with strong preattentively discriminable stimulus differences such as the slopes of lines. A pattern of tilted ellipses in a pattern of upright ellipses gives good grouping (display G in Figure 6) because

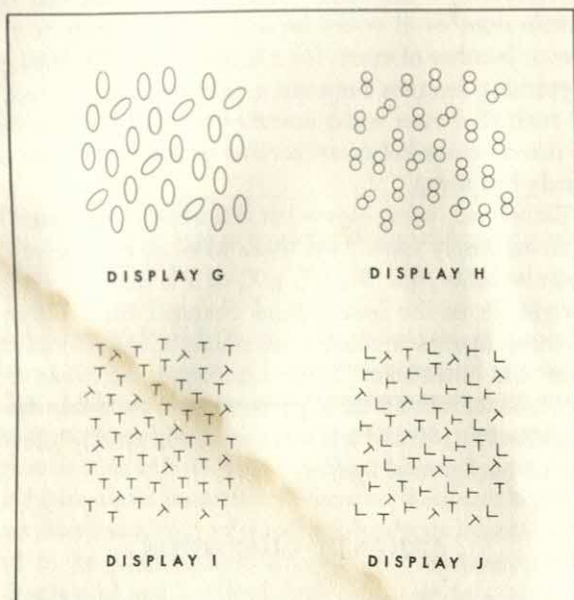


Fig. 6. Displays showing the effect of direction of figures (displays G and H) and of extraneous stimuli (displays I and J) on grouping

direction of the figures is indicated by differences in the slopes of the lines of the ellipses. But grouping is weak when figural direction is indicated by a relational property such as the maximum dimension of a figure defined by the end points of the figure. A pattern of vertical and diagonal figures eight, each made up of two equal circles, produces only weak grouping (display H in Figure 6) because the direction of the figures is indicated only by changes in the direction of the maximum dimension of these figures from vertical to 45 deg, while the slopes of the component lines of the upright and diagonal figures eight remain the same.

Why should the guidance of fixations, peripheral discriminability under uncertainty, and grouping be a function of the same variables? Neisser (1967) has suggested two stages of perceiving: a *preattentive parallel stage*, which responds selectively to stimulus variables to which the visual system is highly sensitive; and a *focal attentive process*, which produces a complete identification of a stimulus based on both a more detailed analysis and stored information. It is possible that grouping is an operation involved in preattentive processing and occurs most strongly for those stimulus variables that guide fixation and to which the preattentive process is sensitive. One should be cautious, however, in assuming a special internal mechanism.

Another possibility is that grouping and peripheral discriminability are a function of those figural properties that are discriminable when attention is not concentrated but distributed over a larger spatial area. One meaning of focal attention is that attention is concentrated or intense in a small spatial area. According to such a definition, *prefocal* attention refers to attention distributed over a large area of the visual field. When attention is thus distributed over a large area, the visual system is more sensitive to simple variables like brightness, color, and line slope than to relational properties.

It is interesting that these simple stimulus variables not only give strong similarity grouping but also underlie the primary segregation of a visual field into regions of uniform visual *texture*. Visual texture is a function of variables that in some sense may be described as variables of brightness and slant, such as mean brightness and the homogeneity and regularity of brightness contrasts (Julesz, 1962). Grouping may in fact thus be described as based on the stimulus variables that give rise to differences of texture. Both grouping and the discrimination of visual texture appear to be based on those stimulus properties that are highly discriminable when attention is not focused but distributed over a larger area.

Wertheimer (1923) proposed that similarity grouping is an associative process based upon stimulus similarities. Similarity grouping, however, may also be described as a segregative process based upon stimulus *differences*. The question of whether grouping involves an underlying associative process that functionally interrelates like elements cannot be answered unequivocally. The experiments reported, however, support the view that similarity grouping is based on the response of the visual system to stimulus differences. It may be that similarity grouping is based on the simultaneous pickup of stimulus differences in local spatial regions. Thus, local differences between disparate and background figures

would suffice to produce grouping even though the disparate figures in one part of a pattern differed from those in another part of the pattern.³ The overall grouping of disparate figures in a pattern would be a function of the grouping in local regions throughout the pattern.

Other factors

The working hypothesis was that similarity grouping is based on the discriminability of stimulus differences when attention is distributed over a patterned visual field. The evidence was that simple figural properties such as differences in line slope are more readily discriminated than relational differences such as the arrangement of lines. It is useful to briefly indicate some further factors that jointly affect similarity grouping and the pickup of stimulus differences in a patterned field.

The first of these is *peripheral acuity*, an important factor determining the stimulus differences picked up when attention is distributed over a larger area. Figures in the periphery are blurred because the eye in the periphery tends to become either myopic or hyperopic (Ferre, Rand, and Hardy, 1931). It is possible that the poor grouping of the disparate X figures in a pattern of upright Ts previously reported by Beck (1966a; 1967) occurs because the arrangement of the lines causes those disparate figures to readily lose discriminability with blur. This possibility is supported by the finding that when the Xs were larger, they grouped as well as the tilted Ts.

The procedure of Experiment II with minor modifications was used to assess the grouping of tilted Ts, sideways Ts, upright Xs, and Ls, all on backgrounds of upright Ts. Two separate groups of 18 subjects each indicated the degree to which the background and disparate figures were seen as separate groups on a scale from 0 to 100. The lengths of the two lines making up the figures in the displays were $\frac{3}{16}$ in. for one group of subjects, and $\frac{1}{16}$ in. for the second group. When the lengths of the lines composing the figures were $\frac{3}{16}$ in., the mean ratings of grouping for the tilted T and the X were the same, 70 and 69 respectively. When the lengths of the lines composing the figures were $\frac{1}{16}$ in., the mean rating of grouping of the tilted T was 53 and of the X 37, a significant difference by a Tukey test ($p < .01$). The relative mean ratings of grouping for the sideways T and the L did not change with the length of the lines composing the figures.

The second factor is that when an individual is instructed to attend to a display, his sensitivity to the stimulus differences can be reduced by the presence of *extraneous stimuli* (Mackworth, 1965; Leibowitz and

Appelle, 1969). The procedure of Experiment I was used to assess the grouping of displays in which the background figures varied. In Figure 6, displays I and J show two of the combinations of disparate and background figures used. In both displays, the disparate figures are tilted Ts, but the background in display J is considerably more complicated (three figures instead of one as in display I). The mean time for counting the disparate figures averaged over 19 subjects for display I was 10.28 sec; for display J it was 14.22 sec (Tukey test, $p < .01$). In general, the results indicate that introducing variability into the background figures reduced the grouping of the disparate figures. Similarity grouping, like peripheral discriminability, is thus interfered with by extraneous difference signals.

A third factor has to do with *horizontality and verticality*. Olson and Attneave (1970) have reported that a pattern made up of vertical and horizontal lines gives better grouping than a pattern made up of 45- and 135-deg lines, even though the difference between the slopes of the lines is 90 deg in both instances. Their interpretation was that grouping occurs among descriptors based on the relationships of a stimulus pattern to a Cartesian reference system. An interpretation consistent with the suggestion that grouping is primarily an acuity function is that a vertical-horizontal frame of reference facilitates the discrimination of lines that coincide with the reference system. The procedure of Experiment VIII was used to assess peripheral discriminability in a patterned field. Each target consisted of either all the same lines in the corners of an imaginary square or three lines the same and one line rotated 90 deg. When subjects had to discriminate a horizontal line in a 3-in. square of vertical lines, the mean number of errors was 4.47; when they had to discriminate a 135-deg line in a 3-in. square of 45-deg lines, the mean number of errors was 8 [$t(14) = 5.59$, $p < .001$]. This suggests that the superiority of vertical and horizontal lines in producing similarity grouping and for peripheral discriminability under uncertainty has its origin in a process whereby lines that coincide with the vertical-horizontal frame of reference are more readily discriminated when attention is relaxed and not focused than lines that do not so coincide.

Notes

This research was supported by National Science Foundation Grant GB8027. A preliminary report of this research was read at the meetings of the Psychonomic Society, Saint Louis, Missouri, November, 1969. Revision received for publication March 22, 1971.

1. A Tukey test applied to the untransformed data gave very similar results. The only difference was that the displays with inverted disparate figures (displays C and F) now differed significantly.

2. A further experiment suggests that the relevant variable is the difference in the number of vertical and horizontal lines. The procedure of Experiment I was used to assess the grouping of these four displays: upright and tilted U, upright and sideways U, upright and sideways H, and upright and inverted U. A Tukey test ($p < .01$) applied to the data showed that the first of these displays grouped best. The second- and third-listed displays grouped next in amount, and equally; and the last-named display grouped least. Note that the background and disparate figures in this third display have the same axes of symmetry and differ only in the number of their horizontal and vertical lines.

3. The following example illustrates the point. A E is embedded in a background of short vertical lines; its own vertical bar is made up of horizontal red lines, its top bar of horizontal black lines, its middle bar of horizontal black lines, and its bottom bar of vertical red lines. An associative theory of perceptual grouping would predict that similar figures would be grouped together. For instance, the red horizontal lines of the vertical bar might be expected to group with the red vertical lines of the bottom bar to form a L, or the black horizontal lines of the top and middle bar might be expected to group with the horizontal lines of the vertical bar to form a F. However, an associative theory would not predict that black horizontal lines, red horizontal lines, and red vertical lines would be grouped to form a E. All three types of lines differ from their immediate background, which is made up of black vertical lines, and a difference theory would predict that they should group to form a E. Informal observations indicate that individuals do see a E rather than a F or a L. The author is indebted to Professor Fred Attneave for this example.

References

- Beck, J. 1966a. Slant and shape variables in perceptual grouping. *Science* 154:538-540.
- Beck, J. 1966b. Effect of orientation and of shape similarity on perceptual grouping. *Perception and Psychophysics* 1:300-302.
- Beck, J. 1967. Perceptual grouping produced by line figures. *Perception and Psychophysics* 2:491-495.
- Eichelman, W. H. 1970. Changes in the relative discriminability of slant and configuration differences. Unpublished doctoral dissertation, University of Oregon.
- Ferre, C. E., Rand, G., and Hardy, C. 1931. Refraction of the peripheral visual field. *Archives of Ophthalmology* 5:717-731.
- Hochberg, J. 1965. Toward a general theory of graphic communication: I. The component perceptuomotor skills at different literacy levels. *Project Literacy Reports* (Cornell University) 5:34-47.
- Julesz, B. 1962. Visual pattern discrimination. *IRE Transactions of the Professional Group on Information Theory* 2:84-92.
- Leibowitz, H. W., and Appelle, S. 1969. The effect of a central task on lumi-

nance thresholds for peripherally presented stimuli. *Human Factors* 11: 387-392.

Mackworth, N. H. 1965. Visual noise causes tunnel vision. *Psychonomic Science* 3:67-68.

Neisser, U. 1967. *Cognitive psychology*. New York: Appleton-Century-Crofts.

Olson, R. K., and Attneave, F. 1970. What variables produce similarity grouping? *American Journal of Psychology* 83:1-21.

Wertheimer, M. 1923. Untersuchungen zur Lehre von der Gestalt, II. *Psychologische Forschung* 4:301-350.

Williams, L. G. 1967. The effect of target specification on objects fixated during visual search. In *Attention and Performance*, ed. A. F. Sanders, pp. 355-360. Amsterdam: North Holland.

continued from page 155

- Atkinson, Richard C., introd. *Contemporary Psychology: Readings from "Scientific American."* San Francisco: W. H. Freeman, 1971. Pp. viii, 484. \$12.00; paperback, \$5.95.
- Balakian, Anna. *Andre Breton, Magus of Surrealism.* New York: Oxford University Press, 1971. Pp. xii, 289. \$10.00.
- Ban, Thomas A., and Lehmann, Heinz E. *Experimental Approaches to Psychiatric Diagnosis: Psychometric Conditioning and Psychopharmacological Studies.* Springfield, Ill.: Charles C Thomas, 1971. Pp. ix, 119. \$10.75.
- Bardwick, Judith M. *The Psychology of Women: A Study of Bio-Cultural Conflicts.* New York: Harper & Row, 1971. Pp. vii, 242. \$7.95; paperback, \$3.95.
- Bell, James Edward. *A Guide to Library Research in Psychology.* Dubuque, Iowa: Wm. C. Brown, 1971. Pp. xii, 211. Paperback, \$2.95.
- Blodgett, Harriet E. *Mentally Retarded Children: What Parents and Others Should Know.* Minneapolis: University of Minnesota Press, 1971. Pp. 165. \$5.95.
- Bott, Elizabeth. *Family and Social Network.* 2d ed. New York: Barnes & Noble, 1971. Pp. xxx, 363. \$10.50.
- Bourne, Lyle E., Jr.; Ekstrand, Bruce R.; and Dominowski, Roger L. *The Psychology of Thinking.* Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xi, 382. \$9.95.
- Browning, Robert M., and Stover, Donald O. *Behavior Modification in Child Treatment.* Chicago: Aldine, 1971. Pp. vii, 422. \$11.95.
- Brush, Robert F., ed. *Aversive Conditioning and Learning.* New York: Academic Press, 1971. Pp. xii, 626. \$25.00.
- Carney, Richard E., ed. *Risk-taking Behavior.* Springfield, Ill.: Charles C Thomas, 1971. Pp. xii, 211. \$13.75.
- Caro, Francis G., ed. *Readings in Evaluation Research.* New York: Russell Sage Foundation, 1971. Pp. xiv, 418. \$12.50.
- Christman, R. J. *Sensory Experience.* Scranton: Intext Educational Publishers, 1971. Pp. vii, 408.
- Clark, David H. *Administrative Therapy: The Role of the Doctor in the Therapeutic Community.* New York: Barnes & Noble, 1971. Pp. xi, 160. Paperback, \$2.50.
- Cleugh, M. F. *Discipline and Morale in School and College: A Study of Group Feeling.* New York: Barnes & Noble, 1971. Pp. ix, 173. \$5.50.
- Colman, Arthur D. *The Planned Environment in Psychiatric Treatment: A Manual for Ward Design.* Springfield, Ill.: Charles C Thomas, 1971. Pp. xvii, 147. \$9.00.
- Coombs, Robert H., and Vincent, Clark E., eds. *Psychological Aspects of Medical Training.* Springfield, Ill.: Charles C Thomas, 1971. Pp. xvii, 566. \$26.00.
- Corcoran, D. W. J. *Pattern Recognition.* Baltimore: Penguin Books, 1971. Pp. 223. Paperback, \$2.45.
- Cottle, Thomas J. *Time's Children: Impressions of Youth.* Boston: Little, Brown and Co., 1971. Pp. xxviii, 354. Paperback, \$3.95.

continued on page 56

RETENTION OF SENTENCES, ANOMALOUS SEQUENCES, AND RANDOM SEQUENCES

William Epstein

University of Wisconsin

Pairs of two-digit stimuli and three-word responses (sentences, S; anomalous, A, or random, R, sequences) were learned to a criterion of % correct; retention was tested immediately, 5 and 24 hours and 1 week later. The lists were acquired at different rates ($S > A > R$) and several measures of retention showed similar differences, but there was no evidence of differential forgetting of the three types of material.

There is ample evidence that linguistic features influence acquisition and immediate memory (Johnson, 1968a; Kintsch, 1970, ch. 8). The present experiment was designed to provide data on the influence of linguistic variables on long-term retention. For this purpose, we compared the retention of short sentences, anomalous sequences, and random sequences at various intervals up to one week. The experimental data are relevant to a variety of considerations, which will be explored in the concluding discussion.

METHOD

—Materials—The learning materials are reproduced in Table 1. There were three lists of eight pairs. In each list the stimuli consisted of the same eight two-digit numbers. The responses were three-word sets (triads) adapted from Marks and Miller (1964) and Epstein (1969). In list S the triads were sentences. In list A, the triads were anomalous sets—syntactically acceptable but semantically anomalous. In list R the triads were random sets. The same 24 words were used in each of the three lists, with the one exception that “smell” replaced “ring” in list A.

—Subjects—The subjects were 108 students from an introductory psychology course. They were assigned to the 12 experimental conditions in nine randomized blocks of 12 subjects each.

—Procedure—The experiment was a 4×3 factorial design: four retention intervals (0, 5, and 24 hr, and 1 wk) and three types of material (S, A, R). The procedure during learning was identical for all 12 conditions. The list was presented in a memory drum at a 4-sec rate, with 8 sec between trials. Learn-

Table 1. Learning materials

List stimuli	List responses		
	Sentences (S)	Anomalous (A)	Random (R)
33	morning stars twinkle	morning boys bark	morning sing dogs
45	wild horses gallop	wild chairs smell	wild hurt trees
16	furry dogs bark	furry bells hurt	boys twinkle furry
75	grateful boys sing	grateful stars bend	papers grateful ring
69	tall trees bend	tall papers twinkle	horses tall bark
98	church bells ring	church trees gallop	church tear chairs
20	pink papers tear	pink horses sing	bend bells pink
13	hard chairs hurt	hard dogs tear	stars hard gallop

ing was by the method of paced anticipation. The list was presented to each subject in four different orders, and learning continued to the criterion of six out of eight correct responses. The subjects were told that for a response to be counted as a correct anticipation, the three words in the triad had to be given in correct order.

Immediately after the criterial trial, a recall test was administered to subjects in the three groups assigned to retention interval 0. The stimuli were presented in the memory drum at a 4-sec rate and the subjects provided the responses. The subjects in the other nine groups left the laboratory after list learning, and they returned either 5 hours, 24 hours, or 1 week later, at which time the recall test was administered. In subsequent references, designation of these conditions are abbreviated as in the examples S-0, S-24, A-5, R-wk.

Another feature of the procedure is worth noting. When the subjects volunteered for participation, they were informed that the experiment was designed to determine the correlation between performances on a verbal-learning task and a visual-discrimination task. An experimental arrangement for securing brightness discriminations was present in the laboratory room. The subjects in the 5-hour, 24-hour, and 1-week conditions were told that the visual-discrimination test would be administered at the second appointment. The retention test was not mentioned. However, at the time of the retention test, instructions were administered to reestablish the task set and to motivate the subjects to recall as many responses as possible.

RESULTS

Acquisition

Table 2 presents the mean number of trials required to attain the criterion of six correct responses. Examination of Table 2 shows that there were no differences among the four sets of groups assigned to the four retention intervals. However, as expected, there were differences among the three lists. List S was acquired most rapidly, list R least rapidly, and

Table 2. Mean number of trials to criterion of six correct responses

List	Retention interval									
	0		5 hours		24 hours		1 week		Combined	
	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
S	8.78	3.00	6.33	2.01	6.11	2.38	6.78	2.39	7.00	2.69
A	10.56	3.58	14.11	4.87	13.33	4.11	12.78	4.48	12.69	4.74
R	15.00	5.53	12.11	4.44	15.89	4.06	15.78	4.40	14.69	5.21
Combined	11.45	5.32	10.85	5.21	11.78	5.76	11.81	5.33		

Table 3. Percent forgetting based on recall scores on immediate (0) retention test

List	Percent forgetting: Words			Percent forgetting: Triads			
	5 hr	24 hr	Combined	5 hr	24 hr	1 wk	Combined
S	24%	28%	64%	25%	28%	66%	40%
A	50	27	76	54	29	80	54
R	41	30	67	60	41	84	62
Combined	38	28	69	46	33	77	

list A was intermediate. The inversion of the acquisition rates for conditions A-5 and R-5 was due to two highly atypical slow learners in condition A-5. An analysis of variance of the trials-to-criterion data yielded only a significant main effect for lists [$F(2, 96) = 29.77, p < .001$]. The main effect of retention interval ($F < 1$) and the interaction effect ($F = 1.49$) did not approach significance.

Three additional measures of acquisition were computed. First, the mean number of triads correctly anticipated on the criterial trial ranged from 6.11 to 6.78. For the three groups assigned to the immediate retention test, the means were 6.78 ($SD = .90$), 6.11 ($SD = .33$) and 6.11 ($SD = .33$) for conditions S-0, A-0, and R-0 respectively. Second, the mean number of triads correct, disregarding order of output within the triad, on the criterial trial ranged from 6.33 to 6.78. For the three groups assigned to the immediate retention test, the means were 6.78 ($SD = .90$), 6.11 ($SD = .33$) and 6.44 ($SD = .73$) for conditions S-0, A-0, and R-0 respectively. Third, the mean number of words correctly anticipated on the criterial trial ranged from 19.33 to 20.78. For the three groups assigned to the immediate retention test, the means were 20.78 ($SD = 2.47$), 19.77 ($SD = 1.22$), and 20.22 ($SD = 2.12$) for conditions S-0, A-0, and R-0 respectively.

Retention

Figure 1 shows the mean total number of words recalled, the first measure of retention. This total includes cases of partial recall as well as cases of complete triad recall, the only restriction being that the recalled item be given as a response to the appropriate stimulus. The retention scores on the immediate test exhibited the typical relationship between rate of acquisition and magnitude of postcriterial drop. The difference was .12, .34, and 1.34 items for conditions S-0, A-0, and R-0 respectively. Analysis of variance of the data summarized in Figure 1 showed a significant main effect of retention interval [$F(3, 96) = 34.67, p < .001$]. The effect of lists ($F = 3.80$) and the interaction effect ($F < 1$) were not significant.

The left half of Table 3 shows the percent forgetting (loss scores) based on the recall scores for the immediate retention test. The average percent forgetting at 5 hours was 38%, the 24-hour average showed a decline, but at 1 week forgetting had increased to 69%. The differences among lists were relatively small and are confined largely to the test at 5 hours.

The second measure of retention was triad recall. The left half of

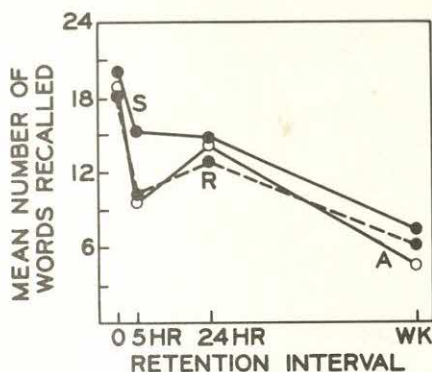


Fig. 1. Mean number of individual words recalled

Figure 2 shows the mean number of complete triads recalled (T_s). Correct recall was credited only if all three words in the triad were given in the correct order to the appropriate stimulus. Again the retention scores on the immediate test exhibited the usual relationship between rate of acquisition and magnitude of postcriterial drop. The difference was .12, .77, and 1.56 for conditions S-0, A-0, and R-0 respectively. Analysis of variance of the data summarized in Figure 2 showed significant main effects for both lists [$F(2, 96) = 16.64, p < .001$] and retention interval [$F(3, 96) = 36.30, p < .001$]. The interaction effect was not significant ($F < 1$). Recall of triads diminished as the retention interval increased, and at each interval the rank order of recall was $S > A > R$.

The percent forgetting is shown in the right half of Table 3. The trend over time is similar to the trend obtained for number of words recalled (left side of Table 3). One additional feature of Table 3 warrants attention. A comparison of the percentages for the two measures of recall shows that whereas the percentages were highly similar for lists S and A, this was not so for list R. In this latter case, comparisons at each of the three retention intervals reveal higher loss scores for recall of triads.

A supplementary measure of triad recall was obtained using a more liberal criterion of correctness (T_L). A triad was scored correct if all three words were recalled, regardless of order. The T_L means are also plotted in Figure 2, whose two parts show that the stringent and liberal scoring criteria yielded identical results for list S and list A. But a consistent difference between T_s and T_L was obtained for list R. The T_L scores were consistently higher, with the result that lists A and R did not differ on this measure. Analysis of variance of the T_L data showed significant main effects of lists [$F(2, 96) = 7.95, p < .01$] and retention interval

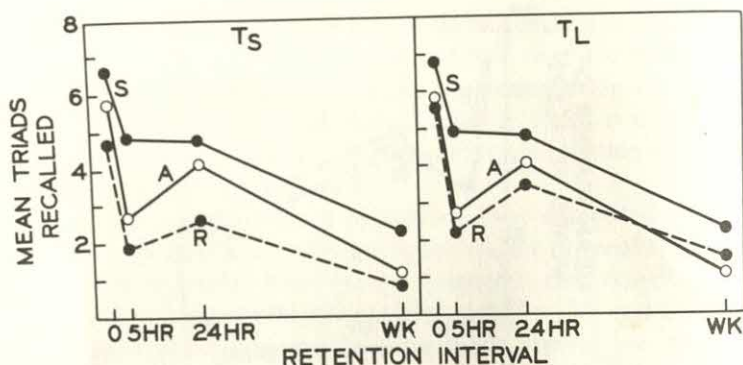


Fig. 2. Mean number of triads recalled correctly, T_S ; mean number of triads recalled correctly disregarding order, T_L

[$F(3, 96) = 32.93, p < .001$]. Individual comparisons among the lists, using the t test, showed significant differences at each interval between list S and the other lists ($.10 > p < .01$). The differences between lists A and R were not significant.

The final retention measure was a conditional probability. For each experimental condition we computed the probability of correct triad recall by the stringent criterion, given that at least one word from the triad was correctly recalled. On the criterial trial, $p(T_S/W)$ ranged from .86 to 1.00. For the three groups assigned to the immediate retention test, $p(T_S/W)$ on the criterial trial was .97, .86, and .87 for conditions S-0, A-0, and R-0 respectively. Comparing these values with those obtained on the immediate recall test showed postcriterial drops of .01, .00, and .14 for conditions S-0, A-0, and R-0 respectively.

The $p(T_S/W)$ measure is plotted at the left of Figure 3. There were consistent differences among the lists; at each interval $S > A > R$. A value for $p(T_S/W)$ was determined for each subject, and an analysis of variance was computed for these data. The analysis showed significant main effects of retention interval [$F(3, 96) = 13.66, p > .001$] and lists [$F(2, 96) = 25.17, p > .001$]. The interaction of lists and retention interval [$F(6, 96) = 1.17$] was not significant.

The $p(T_S/W)$ differences among the lists were due to differences in T_S . The opportunity for triad recall, as defined by the frequency of triads from which at least one word was recalled (W), did not differ for the three lists at each interval. For example, considering lists S and R, the W frequencies at the four retention intervals were S-0 = 63, R-0 = 59; S-5 = 49, R-5 = 49; S-24 = 45, R-24 = 51; S-wk = 24, R-wk = 25.

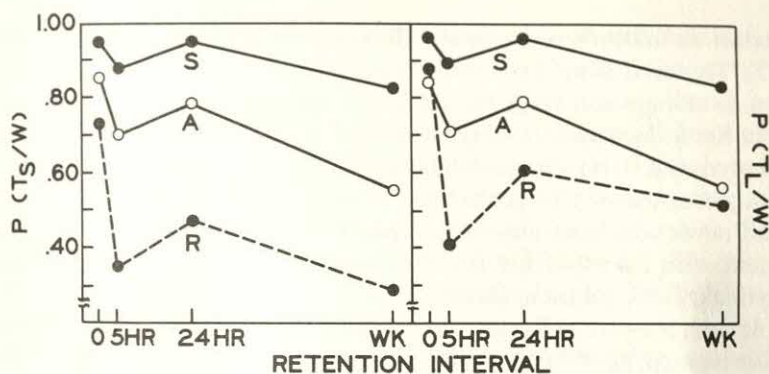


Fig. 3. Conditional probability of recall of complete triad given that at least one word from the triad is correctly recalled, $p(T_s/W)$; conditional probability of recall of complete triad, disregarding order, given that at least one word from the triad is correctly recalled, $p(T_L/W)$

Figure 3 also plots the conditional probabilities using the liberal criterion for scoring triad recall, $p(T_L/W)$. The effect of the liberal score was to raise the values for list R, so that the values of $p(T_L/W)$ for lists A and R become more similar.

DISCUSSION

Five main findings on the retention of S, A, and R triads have been reported. First, the total number of words recalled from lists S, A, and R at each interval did not differ. Second, the number of triads correctly recalled (T_s) at each interval was in the order $S > A > R$. Third, at each interval the conditional probability $p(T_s/W)$ differed for the three lists, $S > A > R$. For list S, $p(T_s/W)$ remained at a very high level even after one week. Fourth, the number of triads recalled disregarding order (T_L) was greatest for list S, but lists A and R were equivalent on this measure. Fifth, with the liberal contingency measure, $p(T_L/W)$, list S again yielded high values at each interval, list A was also unaffected, but list R values were increased relative to the $p(T_s/W)$ measure, so that list A and list R did not differ consistently.

Unitization

The three lists were acquired at different rates. These differences may be attributed to the facilitating effect of semantic and syntactic con-

straints on unitization or response integration (Johnson, 1968a, pp. 435-445). The conditional probabilities provide further evidence of unitization in storage and retrieval. Although the recall of words and triads from list S decreased sharply over time, $p(T/W)$ remained almost unchanged and never dropped below .84. In other words, if at least one item was retrieved, the probability of retrieving all three members of a triad never was lower than .84, even after a week. This finding is consistent with the view that the triads in list S are units (Horowitz and Prytulak, 1969; Johnson, 1968b).

At each interval, $p(T_s/W)$ was higher for list A than for list R. On the other hand, $p(T_L/W)$ was not consistently different for the two lists. The latter finding suggests that syntactical constraint does not facilitate item retrieval. However, syntax can facilitate the retrieval or reconstruction of order, leading to higher values of T_s for list A than for list R. And given equivalent W values for the two lists, $p(T_s/W)$ will be higher for list A than for list R.

Semantics and syntax

The triads in list S combined semantic and syntactic constraint, list A incorporated only syntactic constraint, and the triads in list R were free of any constraints. The experimental results show that these constraints enhanced recall. Semantic constraints act to facilitate retrieval of items, and syntactic constraints act to determine order of output. This conclusion is suggested by the fact that recall of list S was superior on both measures of triad recall, whereas list A was superior to list R only for recall of triads in correct order. The superiority of list S reflects the facilitation of item retrieval due to semantic constraint. The absence of a difference between lists A and R on T_L implies that syntax does not contribute to item retrieval. The superiority of list A on T_s was due to the loss of order information for list R, suggesting that the presence of syntax in list A facilitates the correct order of output.

Forgetting

There was no evidence of differential forgetting among the list types. None of the retention measures yielded a significant interaction of list and retention interval. This is a disappointing outcome. On the grounds of interference theory (Keppel, 1968), one would expect differential forgetting to stem from differences in the magnitude and type of interference exerted by normal language behavior outside the laboratory. The normal sequence of *form classes* (syntax) would interfere with the un-

grammatical sequences of classes in the triads of list R and thus lead to loss of order information. The normal sequence of *words* in meaningful sentences would interfere with triads from both lists A and R and thus lead to loss of items from these lists. Both sources of interference would operate to produce forgetting of list R, whereas only word-sequence interference would affect list A. Word-sequence interference would probably also be present for list S, but to a lesser degree inasmuch as these sequences are less likely to conflict with normal language. This line of reasoning predicts differential forgetting of the lists in the order $R > A > S$. The data did not confirm that prediction.

What accounts for the absence of differential forgetting? One possibility is that the linguistic dimensions that distinguished the lists were not the dimensions along which interference develops. This seems unlikely, and there is in fact contrary evidence in the form of Heisey and Duncan's (1971) demonstration of the buildup and recovery from proactive inhibition as a function of variations of sentence structure in Wickens' (1970) paradigm. Of course, there is an important difference between the conditions of those experiments and the present one. The former were studies of intralaboratory interference whereas the one we report was a study of interference presumed to arise from extralaboratory verbal behavior. In this respect, our experiment is more appropriately compared to other studies of extraexperimental interference in relation to the nature of the material—studies whose results show an unhappy concordance with our data (see Underwood and Ekstrand, 1966). In both cases, the amount forgotten has been independent of the nature of the material and has varied only as a function of degree of learning. A hypothesis worth considering is that the difference between the contexts of verbal behavior in the laboratory and language behavior in extralaboratory discourse is so great that the two domains of behavior are effectively isolated and, therefore, insulated from interfering with each other.

Notes

Received for publication April 16, 1971.

References

- Epstein, W. 1969. Recall of word lists following learning of sentences and of anomalous and random strings. *Journal of Verbal Learning and Verbal Behavior* 8:20-25.

- Heisey, J. A., and Duncan, C. P. 1971. Syntactical encoding in short-term memory. *Journal of Verbal Learning and Verbal Behavior* 10:95-100.
- Horowitz, L. M., and Prytulak, L. S. 1969. Redintegrative memory. *Psychological Review* 76:519-531.
- Johnson, N. F. 1968a. Sequential verbal behavior. In *Verbal behavior and general behavior theory*, ed. T. R. Dixon and D. L. Horton. New York: Prentice-Hall.
- Johnson, N. F. 1968b. The influence of grammatical units on learning. *Journal of Verbal Learning and Verbal Behavior* 7:236-240.
- Keppel, G. 1968. Retroactive and proactive inhibition. In *Verbal behavior and general behavior theory*, ed. T. R. Dixon and D. L. Horton. New York: Prentice-Hall.
- Kintsch, W. 1970. *Learning, memory and conceptual processes*. New York: Wiley.
- Marks, L., and Miller, G. 1964. The role of semantic and syntactic constraints in the memorization of English sentences. *Journal of Verbal Learning and Verbal Behavior* 3:1-5.
- Underwood, B. J., and Ekstrand, B. R. 1966. An analysis of some shortcomings in the interference theory of forgetting. *Psychological Review* 73: 540-549.
- Wickens, D. D. 1970. Encoding categories of words: An empirical approach to meaning. *Psychological Review* 77:1-15.

IMAGES AS MEMORY AIDS: IS BIZARRENESS HELPFUL?

Stanley C. Collyer, John Jonides, and William Bevan
Johns Hopkins University

Eighty subjects learned noun-verb-noun triplets suggesting either common, plausible scenes or bizarre, implausible scenes and under instructions to use imagery or a verbal-memorization strategy. Clearly the best free-recall performance was by the subjects who used plausible images as mediators, a superiority that was not due simply to a difference in semantic meaningfulness between the common and bizarre triplets.

The role of mediating images as aids to memory is currently enjoying a resurgence of interest, as may be seen in the recent reviews of Paivio (1969) and Bugelski (1970). A number of experiments have shown that when subjects are given instructions to use images as mediators in a paired-associates situation, recall scores are remarkably high. For example, the frequently cited study of Wallace, Turner, and Perkins (1957) reported approximately 95% correct recall with 700 paired associates, each seen only one time. More recently, Bugelski, Kidd, and Segmen (1968) also reported that, with enough time, subjects given instructions to form images outperformed controls given no such instructions, a result confirmed by Bugelski (1968) using a more difficult task. Similarly, Paivio (1969) has repeatedly shown image instructions to be superior to rote-repetition instructions or to no instructions at all. Furthermore, the assertion that subjects, when instructed to, do indeed use images instead of verbal mediators was supported indirectly by Bugelski (1968), who examined his subjects' descriptions of their mediators, and in a different fashion by Atwood (1969), who showed that visual interfering stimuli were more detrimental than auditory ones when subjects were instructed to use images, while the reverse was the case in a verbal-learning task not involving imagery. It thus would seem to be well established that images can be used effectively as mediators for learning verbal material.

Attention has now turned to specifying the characteristics of a 'good' mnemonic image—to those qualities that make it most effective as a mediator. Atwood (1969), for example, has demonstrated the importance

of *figural unity* in images. The interdependence of the various parts of the image enhances the likelihood that one part will be able to evoke the whole.

One qualitative aspect of images traditionally thought to be related to mnemonic effectiveness is their *bizarreness*, or implausibility. Advice to construct bizarre and unusual mental pictures can be found in the classical literature (see Yates, 1966), as well as in modern sources (Miller, Galanter, and Pribram, 1960). However, current attempts to confirm bizarreness as an important variable have so far produced inconclusive results (Atwood, 1969; Wood, 1967; Bugelski, 1970), probably due in part to the difficult methodological problems involved. One of these problems involves manipulating the degree of bizarreness of the images without destroying any relevance the mnemonic exercise might have for a real situation in which the subject is trying to remember a number of things. On the one hand, when the experimenter completely specifies for the subject the details of the mnemonic image, the situation is quite different from the latter's being left free to create his own. Although Briggs, Hawkins, and Crovitz (1970) did obtain high recall scores when images were specified, performance might well have been better if the subjects had been required to construct images of their own choosing. On the other hand, simply telling the subjects to construct either common or bizarre images leads to difficult description and rating problems. Furthermore, it probably also confounds figural unity with bizarreness, since the most 'common' image involving two or more initially unrelated objects would likely be a simple juxtaposition.

The experiment now reported sought to clarify the relationship between bizarreness of the mnemonic image and efficiency of free recall by following a procedural middle course. Our aim was to maintain high figural unity and, at the same time, to provide just enough guidance in the construction of the image to ensure that degree of bizarreness was manipulated even as the subject was free to construct an image mostly of his own choosing.

METHOD

—Experimental approach—The procedure involved presenting noun-verb-noun triplets with the nouns held constant across groups; the verbs differed to guide the formation of either common or bizarre images. The subject's task was to recall as many of the second nouns as possible within a given period of time. The basic approach, then, was to use a simple two-group between-subjects design. However, certain additional controls were required. While elements of the triplets might be equated for frequency of usage, associabil-

ity, or some similar criterion, the semantic meaningfulness of each triplet quite probably depends on its composition as a whole. Consequently, ease of recall of the second nouns might well vary as a function of the meaningfulness of the verbal combinations used, variations which could obscure differential recall scores due to imagery. Accordingly, it was deemed necessary to replicate the simple two-group design with another two groups given instructions to make use of a verbal-mediation strategy instead of an imagery strategy in memorizing the material. The influence of bizarreness in *image-mediated* recall is thereby indicated if the difference in performance between the two imagery groups is reliably larger than the difference for the two verbal groups.

—Subjects—In all, 140 Johns Hopkins undergraduates participated as subjects. Sixty of these were used in two preliminary exercises pertaining to stimulus preparation; the remaining 80, separated randomly into groups of 20, constituted the four cells of the 2×2 between-groups design. All were volunteers paid for their services. Each was tested individually in a 30-min session.

—Stimuli—Each of the four experimental groups received a set of 20 noun-verb-noun triplets. For two groups, pairs of nouns were connected by a verb which suggested a 'common,' or plausible, relationship; for the remaining two, the same nouns were connected by a verb suggesting a 'bizarre,' or implausible, relationship.

The triplets were constructed by first selecting 44 nouns with high imagery and concreteness ratings from the Paivio, Yuille, and Madigan list (1968). On a seven-point scale, mean imagery was 6.61 and mean concreteness 6.89. Next, for each noun a common transitive verb was chosen which was *appropriate* to that noun, that is, one which naturally or frequently occurs in conjunction with the noun. Thus, for example, the verb PULL was paired with the stimulus noun HORSE, BLAST with SHOTGUN, MEASURE with TABLESPOON, and so on. Finally, 44 response nouns (mean imagery, 6.56; mean concreteness, 6.88) were selected to complete the triplets. These nouns were chosen so that they were not, in the judgment of the experimenters, highly frequent or stereotyped associates of the stimulus noun and verb.¹ At the same time, they suggested a relationship that might reasonably occur in real life.² Examples are HORSE PULL ROCK, SHOTGUN BLAST BUILDING, and TABLESPOON MEASURE LEMONADE.

When the 44 common triplets had been constructed, they were divided at random into two sets and the verbs interchanged between the sets, yielding two sets of bizarre triplets with 22 triplets in each set. Examples are HORSE AWAKEN ROCK, SHOTGUN MEASURE BUILDING, and TABLESPOON BLAST LEMONADE.

Next, the total of 88 triplets was rated for bizarreness by two groups of ten judges, each group rating 22 common and 22 bizarre triplets. This rating was necessary not only to verify that bizarreness had been manipulated in the desired fashion but also to permit the elimination of triplets for which there was no general agreement or for which subjects could not construct an image. The term 'bizarre' was defined for the judges by presenting such synonyms as 'grotesque,' 'implausible,' 'incongruous,' 'ludicrous,' 'odd,' and 'strange.' Each judge was asked to imagine a scene suggested by each triplet and to rate his image on a seven-point scale from least to most bizarre. He was cau-

tioned not to rate his difficulty in constructing an image but rather its bizarreness once constructed. As a result of these ratings, 8 triplets were discarded, and four groups of 20 remained, two composed of common and two of bizarre triplets. The mean ratings for these groups were: $\text{Common}_1 = 1.70$, $\text{Common}_2 = 1.71$; $\text{Bizarre}_1 = 4.83$, $\text{Bizarre}_2 = 4.82$.

Tests of free recall were then conducted for these four groups (with ten subjects per group), using the procedure described below for the imagery groups. As anticipated, the means were identical across lists employing different verbs [$\text{Common}_1 + \text{Bizarre}_2 = \text{Common}_2 + \text{Bizarre}_1$, $t(38) = .152$, *n.s.*]. As a result, Common_2 and Bizarre_2 were discarded; the main experiment proceeded with only two stimulus lists.

—Procedure—Each subject in the two *imagery groups* was told that he was in an experiment on the use of images as aids to memory. He was informed that he would receive a booklet containing a series of noun-verb-noun triplets, one triplet on each page, and that he would later be asked to recall the second noun of each triplet. His initial task was described as constructing a 'mental picture' of a scene in which the first noun did what the verb specified to the second noun (an example was given, along with a description of an image meeting this requirement). He was instructed to concentrate on each triplet until a satisfactory image was constructed, at which time he was to turn the page to the next one. The maximum time permitted for each triplet was 30 sec, with the experimenter recording the time spent on each page.

Immediately after exposure to the 20 triplets, the subject performed a 4-min distraction task, which consisted of searching for targets on a sheet of random letters. Next, a 4-min free-recall period was provided during which he wrote down all the response nouns he could remember, without regard to their order of presentation.

Each subject in the *verbal-mediation groups* was told that he was in an experiment on the use of verbal associations as aids to memory. Suggestions of verbal strategies were given (e.g., constructing a short sentence containing the three elements of the triplets, or repeating the triplets by rote). All other procedural details were identical to those for the imagery groups.

In order to verify that all the subjects had followed their respective instructions in attempting to memorize the material, each subject was interrogated after the recall test. The experimenter repeated each triplet and asked the subject to report what he did to try to memorize that triplet. If he had failed to use the designated strategy on 8 or more of the 20 items, his data were discarded. On this basis, data from 24 subjects were excluded; 17 of them were from the verbal-strategy groups, and 7 from the imagery groups. Additional subjects were run until each of the two instructional groups had 40.

RESULTS

Free recall

Figure 1 presents the mean number of response nouns correctly recalled by the several experimental groups. Inspection reveals that subjects who had received the common triplets performed distinctly better

during free recall than those presented the bizarre triplets [mean recall difference between types of triplets = 3.2 items; $F(1, 76) = 42.60, p < .001$]. Furthermore, subjects who had been instructed to create images from the triplets generally recalled more items than those required to use a verbal strategy [mean recall difference between types of instruction = 2.7 items; $F(1, 76) = 30.30, p < .001$]. Finally, the interaction between stimulus material and instructions was also significant [$F(1, 76) = 9.85, p < .01$].

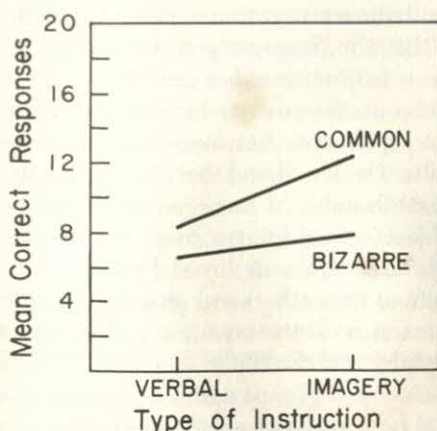


Fig. 1. Mean number of response nouns correctly recalled

The interaction quite clearly was due to the superiority of the common-imagery group. Mean separation tests bear out this conclusion. The mean differences between the group using common images and the other groups were, in order of mean amount recalled, 4.2, 4.7, and 5.9 items respectively [least significant difference ($p < .001$) = 2.36; see Kirk, 1968]. All other differences failed significance. The procedural implications of this interaction are of particular note. Had a significant interaction not obtained, it could be contended that subjects who received the common triplets displayed better free-recall performance solely because these triplets were more semantically meaningful than the bizarre triplets.

Time to construct images

An analysis of the average amount of time spent on each triplet showed no significant differences for the four experimental groups. The mean times ranged from 20.5 sec for the common-imagery group to 23.6 sec

for the bizarre-verbal group. This suggests that the superiority of common images is not simply a consequence of the greater time spent in the construction of those images.

DISCUSSION

The present data indicate that use of common or plausible triplets resulted in superior recall, with unequivocally the best performance being identified with the construction of images from this material. Thus it would appear that the frequently recommended strategy of using bizarre imagery as a mnemonic aid is unwise.

Before speculating on the reasons for this finding, let us explore the possibility that an explanation not involving imagery might easily account for our results. The likelihood that the superiority of the common-imagery group is attributable to some 'verbal' factor rather than to the content of the subjects' visual images was fairly certainly eliminated by the several procedural controls employed. First, the fact that some words are easier to remember than others was offset by using the same stimulus and response nouns in both the common and bizarre triplets. Second, while the meaningfulness of the triplet as a whole changes when the verb is changed, the inclusion of groups with identical stimulus materials, but with instructions to use a verbal-mediational strategy, makes possible the segregation of semantic effects from effects due to imagery. Indeed, the significant interaction of stimulus material and instructions makes clear the operation of nonverbal—here, specifically imaginal—factors in the enhanced recall of the common-triplet response nouns.

How should we account for the superiority of common over bizarre imagery? One might suppose, given the virtually limitless capacity of visual memory (see Wallace et al., 1957; Standing, Conezio, and Haber, 1970), that an image, if well formed, should persist regardless of its bizarreness. Thus it is possible that the subjects' bizarre images for some reason may have been less complete, less detailed, or less 'vivid.' For all groups, the 30-sec maximum for image construction was usually ample; nevertheless, the subjects presented with the bizarre triplets may have adopted a less stringent criterion for vividness or completeness.

A second possibility for the superiority of plausible imagery is that a common triplet may easily suggest several images to a subject, whereas a bizarre triplet suggests only one. A greater number of images leading to a given response could increase the likelihood of that response at the time of recall.

Figural unity, which Atwood (1969) has shown to be important for remembering an image, probably did not play a role in the present experiment. The subjects in both imagery conditions were instructed to create an image that linked the two nouns in some manner specified by the verb, so that there is no reason to believe that the parts of a bizarre image were less spatially interdependent than those of a common image.

One final point concerns what is meant by 'common' as opposed to 'bizarre' images as mnemonic aids. One might object that there is nothing particularly common about ELEPHANT TRAMPLE GARDEN OF QUEEN PLAY HARP (two of the common triplets used); that is, that the triplets do not truly span a continuum from common to bizarre. In fact, 'commonness' is probably an inappropriate term when used in the context of mnemonic imagery. A truly common image in a practical mnemonic system would be an extremely rare event, since the essence of such a system is the pairing of the object to be remembered with one of an arbitrary list of items, a pairing chosen simply because the items are easy to remember in a particular order. It therefore may be more reasonable to refer to a continuum running from 'plausible' to 'bizarre.'

In conclusion, and aside from the particular reasons for the superiority of common over bizarre imagery, it seems fair to say that any practical advice to a person for using imagery as a mnemonic system should, while stressing concreteness and figural unity, direct him to imagine the most plausible connection between the retrieval cue and the object to be remembered, and not to waste effort constructing an especially bizarre, implausible, or unusual image.

Notes

This research was performed under Contract N00014-67-A-0163-0001 between the Johns Hopkins University and the Physiological Psychology Branch, Office of Naval Research. Dr. Collyer's address is now: Technology Incorporated, Life Sciences Division, 8531 North New Braunfels Avenue, San Antonio, Texas 78217. Mr. Jonides's address is now: Department of Psychology, University of Pennsylvania, Philadelphia, Pennsylvania 19104. Received for publication April 23, 1971.

1. Some examples of nouns rejected on this basis would be those in brackets: HORSE PULL [CARRIAGE], SHOTGUN BLAST [THIEF], or TABLESPOON MEASURE [SUGAR].

2. Response nouns rejected on this basis might have been (in brackets) HORSE PULL [STRAWBERRY], SHOTGUN BLAST [FORK], or TABLESPOON MEASURE [HEADLIGHT].

References

- Atwood, G. E. 1969. Experimental studies of mnemonic visualization. Unpublished doctoral dissertation, University of Oregon.
- Briggs, G. G., Hawkins, S., and Crovitz, H. F. 1970. Bizarre images in artificial memory. *Psychonomic Science* 19:353-354.
- Bugelski, B. R. 1968. Images as mediators in one-trial paired associates learning: II. Self-timing in successive lists. *Journal of Experimental Psychology* 77:328-334.
- Bugelski, B. R. 1970. Words and things and images. Presidential address, Eastern Psychological Association, Atlantic City, April.
- Bugelski, B. R., Kidd, E., and Segmen, J. 1968. Image as a mediator in one-trial paired-associate learning. *Journal of Experimental Psychology* 76: 69-73.
- Kirk, R. E. 1968. *Experimental design: Procedures for the behavioral sciences*. Belmont, Calif.: Wadsworth Publishing.
- Miller, G. A., Galanter, E., and Pribram, K. H. 1960. *Plans and the structure of behavior*. New York: Holt.
- Paivio, A. 1969. Mental imagery in associative learning and memory. *Psychological Review* 76:241-263.
- Paivio, A., Yuille, J. C., and Madigan, S. A. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology* 76(1, part 2).
- Standing, L., Conezio, J., and Haber, R. N. 1970. Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science* 19:73-74.
- Wallace, W. H., Turner, S. H., and Perkins, C. C. 1957. Preliminary studies of human information storage. University of Pennsylvania Institute of Cooperative Research, Signal Corps project 1326.
- Wood, G. 1967. Mnemonic systems in recall. *Journal of Educational Psychology* 58(6, part 2).
- Yates, F. A. 1966. *The art of memory*. London: Routledge & Kegan Paul.

INTERFERENCE IN SENTENCE RECALL

Calvin D. Sowder

Chico State College

The 144 subjects learned either one list of sentences or three lists (of varied interlist topical similarity) by the method of free recall. After 24 hours a measure of semantic recall was taken. Significant retroactive ($p < .01$) and proactive ($p < .025$) inhibition were observed as a function of the number of lists. Significant retroactive ($p < .05$), but not proactive, inhibition occurred as a function of interlist topical similarity.

This experiment examines recall of the *content* of sentences learned by the method of free recall. Variables of interest were (a) number of lists and (b) interlist similarity of topics with which the sentences dealt.

The existing data for sentences are equivocal with regard to both of these variables as they affect recall. Hall (1955) failed to find retroactive inhibition in a task requiring recall of missing words from previously presented sentences, while Mehler and Miller (1964) found a negligible amount of retroactive inhibition for the semantic content of sentences learned by the method of free recall. Slamecka (1960, 1961) demonstrated both proactive and retroactive inhibition in the recall of single 20-to-30-word sentences learned by serial anticipation. Slamecka's relearning measure, however, does not permit examination of the recall of meaning without verbatim recall of the sentence.

In the experiment below, topical similarity was manipulated by using lists of sentences that dealt with either biology or Australian history. Number of lists was varied by having groups learn either one or three lists. Recall of semantic content was measured after 24 hours. An interference theory would predict that subjects learning a single list would exhibit greater recall than those learning multiple lists and that subjects recalling a list similar to their other lists would show poorer performance than subjects recalling a dissimilar list.

METHOD

—Subjects—Four groups of 36 subjects (a total of 144) were used. All were students enrolled in introductory psychology at Chico State College and were fulfilling a course requirement.

—Materials—Six lists of 14 sentences each were used. The sentences varied in length from five to nine words. Three lists of sentences dealt with biology and three lists with Australian history. Three random orders of each list were used, and each list was used equally often in each possible position.

—Procedure—During acquisition the sentences were presented with a Kodak Carousel projector. Each list was presented four times with each sentence exposed for 5 sec during each presentation. Exposure time was controlled by an external timer. A 5-min test of written recall followed the second and fourth presentation of each list. The instructions emphasized exact verbatim reproduction of the sentences during acquisition, but scoring of recall was in the 'lenient' fashion described below. Written recall was tested again after 24 hours. The subjects were now given as much time as necessary to recall but were required to take at least 20 min in the three-list conditions and at least 7 min in the single-list condition. They were also told to recall the sentences in any order that occurred to them.

—Design—Three groups each learned three consecutive lists; the fourth group was a single-list control. The similarity relation involved sentences dealing with biology or Australian history. In condition SSS the subjects learned three consecutive lists all dealing with the same topic. In conditions DSS and SSD, the subjects learned two lists dealing with one topic and one list dealing with the other topic. Thus in condition DSS, the first list for half the subjects dealt with biology. Subjects were run in groups of three to six, the groups randomly assigned to treatment orders. Data for extra subjects was randomly eliminated before analysis.

RESULTS

Table 1 presents the mean number of sentences recalled during acquisition and after 24 hours. Sentence recall, both during acquisition and after 24 hours, was scored in a 'lenient' fashion. That is, sentences were counted correct if they had a subject or object the same as or similar to the one presented during acquisition. A second, independent rater scored protocols for 72 of the students and disagreed in only 2 cases. It was therefore assumed that the scoring was reliable. The result of this method was that few sentences were rejected by the scorer and the data tend to reflect total failure by the student-subjects to attempt recall of the meaning of a sentence.

The analysis of data for original learning indicated no significant differences among the groups for the first list (including the control) nor for the third list (including the control), with both $F_s < 1.0$. Thus there is no evidence that the four groups were not equivalent in original learning.

The analysis of recall for the first list after 24 hours used two orthogonal comparisons. The first of these [SSS, SSD, and DSS versus the

Table 1. Means for sentence recall in acquisition and after 24 hours

Group	Acquisition			After 24 hours		
	List 1	List 2	List 3	List 1	List 2	List 3
SSS	10.53	10.80	10.83	7.03	8.64	8.61
SSD	10.11	10.00	10.89	6.89	7.89	9.00
DSS	10.08	10.78	10.61	8.30	8.39	8.56
Control	10.69			10.08		

Note: Each mean is based on 36 observations; the means in acquisition are for the second recall trial of a given list.

single-list control, $F(1, 140) = 27.50, p \leq .01$] indicates significant retroactive inhibition as a function of number of lists. The second comparison [SSS versus DSS, $F(1, 140) = 4.18, p \leq .05$] indicates significant retroactive inhibition as a function of the topical similarity of the intervening lists.

Similarly, the analysis of recall for the third list after 24 hours also used two orthogonal comparisons. The first of these [SSS, SSD, and DSS versus the single-list control, $F(1, 140) = 6.62, p \leq .025$] indicates significant proactive inhibition as a function of number of lists. The second comparison was nonsignificant ($F < 1.0$), indicating that the null hypothesis on the effect of similarity of prior lists was not rejected.

DISCUSSION

These data may be partially accommodated by an interference theory that assumes a process of generalized response competition (Postman, Stark, and Fraser, 1968) under conditions of unlimited recall time like those in the experiment just described. According to this notion, inhibition in recall is proportional to the number and similarity of competing responses. The data only partially support this idea, because lesser inhibition was not demonstrated in the recall of the third list when it was dissimilar to the two previous lists. This failure is not decisive, however, as it may be attributable to the relative degree of similarity, the number of prior lists, and the degree of prior learning (Underwood, 1957).

The lenient method of scoring used here is assumed to reflect a loss of availability of the semantic content of the sentences rather than inaccuracy in verbatim reconstruction. Given this assumption, it may be argued that forgetting of the semantic content of sentences can be explained in the same way as the forgetting observed in studies using traditional paired-associate techniques (Keppel, 1968).

Notes

A version of this paper, entitled "Topical similarity as a variable in the forgetting of sentences learned by the method of free recall," was presented at the meeting of the Western Psychological Association, Los Angeles, California, 1970. Requests for reprints should be sent to Calvin D. Sowder, Department of Psychology, Chico State College, Chico, California 95926. Received for publication June 1, 1971.

References

- Hall, J. F. 1955. Retroactive inhibition in meaningful material. *Journal of Educational Psychology* 46:47-52.
- Keppel, G. 1968. Retroactive and proactive inhibition. In *Verbal behavior and general behavior theory*, ed. T. R. Dixon and D. L. Horton. Englewood Cliffs, N.J.: Prentice-Hall.
- Mehler, J., and Miller, G. A. 1964. Retroactive interference in the recall of simple sentences. *British Journal of Psychology* 55: 295-301.
- Postman, L., Stark, K., and Fraser, J. 1968. Temporal changes in interference. *Journal of Verbal Learning and Verbal Behavior* 7:672-694.
- Slamecka, N. J. 1960. Retroactive inhibition of connected discourse as a function of similarity of topic. *Journal of Experimental Psychology* 60:245-249.
- Slamecka, N. J. 1961. Proactive inhibition of connected discourse. *Journal of Experimental Psychology* 62:295-301.
- Underwood, B. J. 1957. Interference and forgetting. *Psychological Review* 64:49-60.

POLARIZATION IN SERIAL AND PAIRED-ASSOCIATE LEARNING

Richard A. Kasschau
University of South Carolina

The average degree of polarization (D_4) of the words in a list was varied as associative meaningfulness (Noble's m) was held constant. Across three experiments D_4 and number of trials to criterion were significantly positively related, and the effects (among them, the von Restorff isolation effect) that have been found with varying m were essentially replicated with varying D_4 . The exceptions to these findings, and the value of D_4 as a measure of meaningfulness, are discussed.

The influence of meaningfulness on the ease of learning verbal material has been explicitly recognized by experimental psychologists since at least 1885, when Ebbinghaus introduced 'nonsense' syllables. Failing in the attempt to create verbal material devoid of meaning, the effort turned then to quantifying and measuring the relative amounts of meaningfulness possessed by varieties of verbal material. Broadly classified, these many measures may be divided into measures of associative (Noble, 1952; Deese, 1965; and others) and semantic meaningfulness (Osgood, Suci, and Tannenbaum, 1957). Certain difficulties have resulted from this absence of a commonly accepted quantitative measure of meaningfulness—an absence at least partly responsible for the large number of different, although often similar, measures that have been developed. Lacking a single acceptable definition of meaningfulness, it would seem advisable to study the similarities and differences of the semantic and associative measures extant.

The need for such studies is emphasized, first, by the clear disagreement as to the distinction, or lack of it, between the two types of measures. For example, after reviewing Bousfield (1961), Deese says, "Bousfield's results are convincing enough to make it profitable to view the process underlying the semantic differential as an associative one" (1965, p. 72). Osgood, after reviewing the same (and related) work, says, "the reactions of subjects to a semantic differential are not predictable from knowing their word-association hierarchies" (1961, p. 105).

Second, Kasschau and Pollio (1967) have demonstrated that with

associative (or semantic) meaningfulness held constant, words varying in semantic (or associative) meaningfulness are sufficient to mediate response transfer to a second list with a 41%–43% savings. Thus, both semantic and associative measures seem to index aspects of meaningfulness that can influence ease of learning.

Finally, Koen (1962) and Zippel (1967) have reported a Pearson r of $+.76$ between the Noble (1952) association values (m) of an 18-word sample and the average distance from a neutral rating of 4.00 (D_4) of the semantic-differential ratings of those words on 20 scales. By contrast, Howe (1965) reported $r = +.51$ and Paivio (1968) reported $r = +.43$ in similar comparisons. These reports indicate that as the association value of the words increased, so did the rated intensity of meaning on the semantic differential, although there has been some disagreement as to the magnitude of the relationship.

In sum, there is evidence to indicate that semantic and associative measures are correlated but measure distinctive aspects of meaningfulness. To the extent that this correlation exists but falls short of being perfect, there is justification for studying the relation between these measures, especially with respect to the ease of learning verbal material defined by them.

For the words in each of the following experiments, associative meaningfulness had already been shown to yield a consistent change in behavior as a function of increments in m . The procedures of previous studies were replicated insofar as practicable, but m was equated across lists and the materials to be learned were varied in the intensity of rated semantic meaningfulness. Would these variations in D_4 with m held constant yield effects similar to those noted when m was varied?

EXPERIMENT I

Noble (1952) demonstrated that the difficulty in reaching a criterion of one perfect recitation of a 12-item serial list was a decreasing curvilinear function of m . He also showed that for three preexperimental ability levels, the difficulty in attaining a criterion of 7 of 12 items correct on those same serial lists was a decreasing curvilinear function of m , and that in terms of total trials to criterion there was an interaction between the effects of ability and m . Holding m constant, Experiment I was an attempt to replicate these findings using three serial lists of different average D_4 and dividing the subjects in each group into three ability levels based on practice-list performance.

Method

—Subjects—The subjects were student volunteers from the introductory psychology course at the University of South Carolina. Participation in the experiment partially fulfilled a course requirement. There were three groups, composed of 21 females (mean age, 20.8 years) and 42 males (mean age, 19.7 years). Equal numbers of subjects were randomly assigned to each group, keeping the ratio of males to females constant.

—Apparatus and materials—The words used in Experiment I (as in Experiments II and III) were selected from among 96 nouns and adjectives that one group of 30 undergraduates had previously rated on 10 semantic-differential scales and that another 30 undergraduates had used as stimulus words in a 30-sec production test similar to Noble's (1952; see Kasschau, 1970). The 96 words were ranked according to increasing magnitude of D_4 with word 1 having the lowest D_4 . To obtain a high-intensity set of words, words were drawn in sequence starting with word 96 until the requisite number had been obtained. When all lists for an experiment had been selected, it was usually necessary to draw additional words so as to exactly equate each list for number of one- and two-syllable words, replicated initial letters, and average m .

All materials were presented on a Stowe 459B memory drum, in an air-conditioned, sound-proofed, humidity-controlled room.

—Procedure and lists—With the following exceptions, the method described by Noble (1952) was replicated exactly. First, neither rest intervals nor color naming was included in the present procedure. Second, only one practice list was used to establish the preexperimental ability levels, and each subject learned exactly the same list. Third, the practice and experimental lists were both learned in one experimental session, separated only by a 1-min interval during which the experimenter engaged the subject in conversation while rearranging the memory-drum shutters. Finally, 21 subjects served in each experimental group, with 7 at each ability level based on the number of trials to attain a 7/12 criterion on a single 12-item practice list. These levels were: *slow* (it took the subjects more than 12 trials to reach criterion), *average* (8–12 trials to criterion), and *fast* (less than 8 trials to criterion).

The experimental lists were obtained as follows. For the 'low' (low D_4) list, the 12 words drawn from words 1–17 (see section on materials, above) had an average m of 7.00 and an average D_4 of 1.83. For the 'medium' list, the 12 words drawn from words 41–59 had an average m of 7.04 and an average D_4 of 3.59. For the 'high' list, the 12 words drawn from words 77–96 had an average m of 7.34 and an average D_4 of 5.36. The 12 words for the practice list were drawn 6 each from words 18–40 and 60–76 and had an average m of 7.30 and an average D_4 of 3.39.

Results

The performance of the three experimental groups learning the low-, medium-, and high- D_4 lists is plotted in Figure 1, showing the number of trials to reach successive criteria.

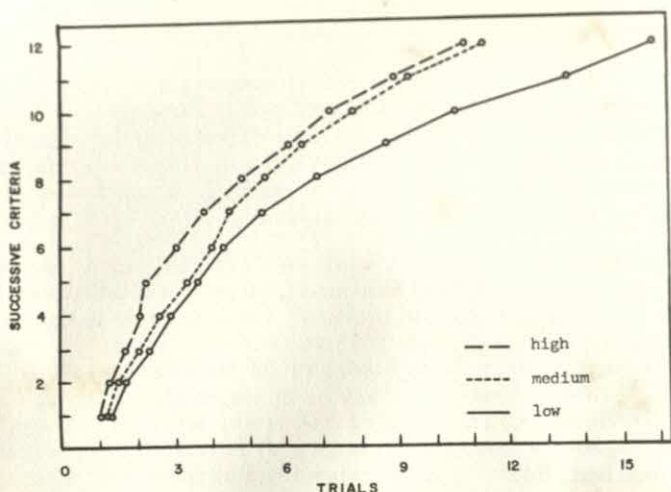


Fig. 1. Mean number of trials to successive criteria, Experiment I

The relationship of difficulty and meaningfulness at successive criteria of correctly anticipated items may be replotted from the data in Figure 1. This relationship was identical in all important respects to that reported by Noble (see Figure 2, 1952) and is not repeated here. The average number of trials to achieve a criterion of one perfect recitation on the experimental list as a function of the three preexperimental ability levels (slow, average, and fast learners) is reported in Figure 2.

As a check on the equality of variances for the groups having differing D_4 lists and ability levels, Hartley's test for homogeneity of variance was calculated [$F_{\max} (df = 6, k = 9) = 3.61$]; it was not significant. Next, a treatment-by-levels analysis of variance was performed on the number of trials to achieve a criterion of one perfect recitation of the experimental list. The main effect of average semantic meaningfulness of the experimental lists was highly significant [$F(2, 54) = 7.02, p < .005$]. The number of trials required to reach criterion in learning the low- D_4 list differed significantly ($p < .005$) from that required to learn either the medium or high lists, which did not differ significantly from one another according to the Duncan multiple-range test. The interaction of treatments and levels was not significant [$F(4, 54) < 1.00$], indicating that the treatments effect was consistent across all levels of preexperimental ability.

As a final comparison, the mean number of errors made in achieving a

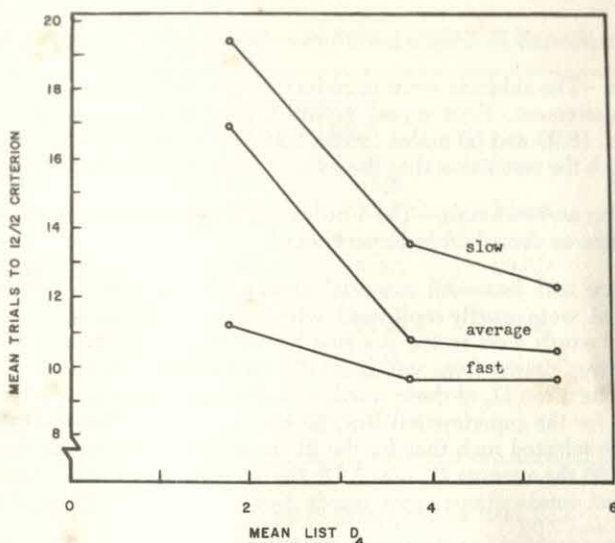


Fig. 2. The difficulty-meaning relationship as a function of number of trials to 7/12 criterion on a practice list, Experiment I

criterion of one perfect recitation as a function of serial position for the low, medium, and high lists was calculated. Except for minor reversals at serial positions 6 and 10-12, where the group working with the medium- D_4 list made slightly fewer errors than the group working with the high list, there was a perfect inverse relationship between errors and average D_4 of the lists.

EXPERIMENT II

Cieutat, Stockwell, and Noble (1958), using variations of high and low m of the stimulus and response terms, demonstrated (a) a greater difference in performance of the groups learning the 'high-high' and 'low-low' lists than any other comparison, (b) that variations in response-term meaningfulness had a significantly greater effect than similar variations in stimulus-term meaningfulness, and (c) that both stimulus- and response-term m as well as the interaction of these two variables in turn interacted with amount of practice. Substituting high and low semantic for associative meaningfulness, Experiment II was an attempt to replicate these findings by varying D_4 while holding m constant.

Method

—Subjects—The subjects were introductory psychology students fulfilling a course requirement. Four equal groups were composed from 24 females (mean age, 18.9) and 56 males (mean age, 19.2). Assignment to groups was random with the restriction that the ratio of males to females be constant.

—Apparatus and materials—The words used were selected and the apparatus was the same as described in Experiment I.

—Procedure and lists—All essential aspects of the method described by Cieutat et al. were exactly replicated, with the following exceptions. First, the 10 pairs of words used as the practice list were one- and two-syllable nouns and adjectives drawn from words 33–62 (see section on materials, Experiment I). The mean D_4 of these words was 3.49; the mean m was 7.025.

Second, for the experimental lists, 40 one- and two-syllable adjectives and nouns were selected such that for the 20 'high' (high D_4) words drawn from words 69–96 the average D_4 was 5.13; the average m was 7.21. Similarly, for the 20 'low' words drawn from words 1–30 the average D_4 was 2.08; the average m was 7.02.

All other significant details of apparatus and procedure were replicated. Thus there were four groups, each learning an identical practice list for 12 trials and then a 'low-low,' 'high-low,' 'low-high,' or 'high-high' experimental list for 12 trials; a correction procedure was used, and both stimulus and response terms were pronounced at a standard 2:2 rate with a 4-sec intertrial interval.

Results

First, a simple randomized-design analysis of variance was performed on each subject's total number of correct anticipations on the practice list. The experimental groups may be considered of equal initial ability, since the obtained $F [(11, 76) = 1.01]$ was clearly nonsignificant.

Table 1 presents the average performance of the four groups in learning the experimental lists. The order of increasing difficulty in terms of the cell mean was 'high-high,' 'high-low,' 'low-high,' and 'low-low.' The average difference in performance due to variations in stimulus-term meaningfulness was greater than that due to variations in response-term meaningfulness.

Figure 3 shows the mean number of correct responses per subject as a function of practice. A tendency toward negative acceleration is apparent in all four curves. All curves appear to be approaching the same asymptote, with the group learning the 'low-low' list generally retaining its inferior position and the group learning the 'high-high' list generally retaining its superior position.

Because all four curves appear to be approaching the same asymptote

Table 1. Mean total correct responses during trials 2-6, Experiment II

D_4 value of stimulus term	D_4 value of response term		Mean	Stimulus difference
	Low (2.08)	High (5.13)		
Low (2.08)	27.05	30.85	29.95	4.53
High (5.13)	32.70	36.25	34.48	
Mean	29.88	33.55		
Response difference	3.67			

and have for all practical purposes overlapped at trial 6, a $2 \times 2 \times 5$ analysis of variance was performed on the subjects' performance over the first five trials, in accord with a recommendation by Anderson (1963). The analysis yielded a significant effect of stimulus meaningfulness [$F(1, 76) = 9.78, p < .001$], a significant effect of response meaningfulness [$F(1, 76) = 4.33, p < .05$], and a significant practice effect [$F(4, 304) = 268.48, p < .001$]. The only other significant effect was the inter-

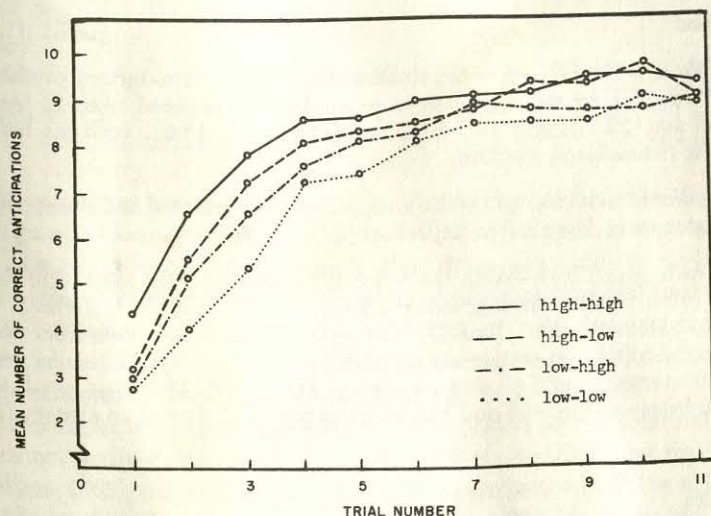


Fig. 3. Mean number of correct anticipations as a function of the number of trials of practice, Experiment II

action of stimulus-term meaningfulness and practice [$F(4, 304) = 2.48$, $p < .05$].

As a final check, Pearson product-moment correlations (r) were calculated between the total correct responses on the practice and experimental lists. The obtained correlations ('high-high' D_4 , .56; 'low-high,' .37; 'high-low,' .61; and 'low-low,' .59) indicated no apparent trend based on ease of learning. The average r (by Fisher's Z transformation) was .54, consistent with the value reported by Cieutat et al.

EXPERIMENT III

Rosen, Richardson, and Saltz (1962) demonstrated that printing in red ink the fifth item of a nine-item serial list of words when all the other items were printed in black resulted in that fifth word being learned more rapidly than without benefit of the red ink. This von Restorff isolation effect was obtained only if average list m was low, not high. In the present experiment m was held constant and the average D_4 was varied across lists. It was predicted that with a low- D_4 list, isolating the fifth item would benefit the ease of learning it, as compared to a nonisolated control, but that isolating it in a high- D_4 list would not aid the learning of it, as compared to a nonisolated control.

Method

—Subjects—The subjects were students fulfilling an introductory psychology requirement. Four equal groups were randomly composed from 12 females (mean age, 18.7 years) and 60 males (mean age, 18.6), with the ratio of males to females kept constant.

—Apparatus and materials—The words used were selected and the apparatus was the same as described in Experiment I.

—Procedure and lists—With the following exceptions, the method of Rosen et al. (1962) was replicated exactly. First, only 72 subjects were used. Second, to obtain the 'low' (low D_4) list, nine words were drawn from among words 1–14 (see the section on materials, Experiment I). These nine words had an average m of 7.37 and an average D_4 of 1.85. The 'high' words were drawn from words 81–96 and had an average m of 7.38 and an average D_4 of 5.50.

Results

For each subject, rank 1 was assigned to the list item eliciting the greatest number of correct anticipations, rank 2 to the item eliciting the

second greatest number, and so forth down to rank 9 for the item eliciting the least number. The mean ranks of the isolated and control items in the high and low lists are presented in Table 2. The main isolation effect was significant [$F(1, 68) = 4.69, p < .05$], indicating that the mean rank of the isolated item was significantly lower than the rank of the same item when it was not isolated. However, the interaction of isolation and meaningfulness failed to achieve significance [$F(1, 68) = 1.06, p < .20$], indicating that the effect of isolation on number of correct anticipations was consistent at both levels of semantic meaningfulness, or D_4 .

Table 2. Mean rank within list and mean correct anticipations of the isolated and control terms for 15 learning trials, Experiment III

List	Ranks		Correct Anticipations	
	Mean	Standard deviation	Mean	Standard deviation
Low D_4 (1.85)				
Isolated	5.78	2.40	10.11	3.25
Control	6.36	2.25	9.50	3.96
High D_4 (5.50)				
Isolated	5.61	1.95	11.06	2.76
Control	7.25	1.82	9.67	2.92

The mean number of correct anticipations to the isolated fifth item for each group is also presented in Table 2. The same general conclusions can be reached for this data, namely, that there was a difference in the mean number of correct anticipations of isolated as opposed to control items, but no interaction between isolation and list meaningfulness. However, as Rosen et al. noted (1962), these results are less tenable due to the potentially more rapid learning of the high list.

A two-factor analysis of variance was performed on the total number of correct anticipations made by each subject. As might be expected (Wallace, 1965), the isolation versus control comparison failed to achieve significance based on this measure [$F(1, 68) < 1.00$]. However, the high versus low comparison was significant [$F(1, 68) = 4.27, p < .05$], indicating that a greater number of correct anticipations was elicited by

the high- D_4 list. The interaction effect was nonsignificant [$F(1, 68 < 1.00)$], indicating consistency of the high versus low difference across both isolated and control lists.

DISCUSSION

Variations in average list D_4 produce very apparent differences in the ease of learning those lists in a diversity of experimental situations where m is held constant. In all the present experiments, high- D_4 lists were learned in fewer trials and with fewer errors than lower- D_4 lists. In each experiment there were, however, some interesting deviations from the results of the previous experiments that used m as the primary independent variable. Those deviations are discussed next.

Experiment I

The original study by Noble (1952) was not replicated in two respects. First, although the overall effect of D_4 variation was significant and the trend of more trials required to learn lower- D_4 lists was consistent, the medium- and high- D_4 lists did not differ significantly in the total average number of trials to reach criterion. It should be noted, however, that the average m of the three lists was 7.13, as measured using Noble's (1952) technique with only a 30-sec interval per word. In terms of the average number of associations, it would seem safe to assume that the words used here were of approximately similar meaningfulness to the most meaningful words used by Noble (e.g. KITCHEN, $m = 9.61$, using a 1-min production interval with more extensive editing of 'acceptable' responses). Thus, the demonstrated effects of D_4 variation were over and above any effects attributable to a substantial degree of associative meaningfulness.

Second, the heterogeneity of the data composing the treatments-by-levels cell means noted by Noble was not obtained. This presumably resulted from the relatively similar number of trials required by each group to reach criterion. Related to this was the consistency across the ability groups' relative numbers of trials to criterion as a function of D_4 . Thus, Noble's hypothesis that "slow learners are more sensitive to differences in meaningfulness than are relatively faster learners" may be limited to differences in associative meaningfulness, since the interaction of D_4 and levels fell far short of significance.

Noble's general conclusion that difficulty in serial learning is a decreasing exponential function of a list's m value was also true of its D_4 value.

Likewise, there was a relatively perfect inverse relationship between the mean number of errors and list D_4 . In sum, the general effectiveness of list- D_4 variations in producing trials-to-criterion variations was well demonstrated in serial learning. In those instances that the results failed to replicate Noble's, the causes can generally be traced to the high average m of the words in the present study.

Experiment II

The finding of Cieutat et al. (1958) that variations in response m yielded greater differences in performance than similar variations in stimulus m was not confirmed for response and stimulus D_4 . In fact, there was a tendency for the reverse to be true. With only minor exceptions (see Figure 3), performance in the group learning the 'high-high' list (both stimulus and response of high D_4) was better than that in the group learning the 'high-low' list, which was better than the group learning the 'low-high' list, which in turn was better than the group learning the 'low-low' list. This order of performance is more consistent with Paivio's (1968) observation that the superior effect of response m has been obtained only with nonsense words as low- m items and that m may be more potent on the stimulus side when varied entirely among familiar words. The present results extend this hypothesis to variations of D_4 among familiar words. Although the present words were originally selected without respect to imagery (Kasschau, 1970; Paivio, 1968), an after-the-fact analysis of those words for their imagery ratings by Spreen and Schulz (1966) or by Paivio, Yuille, and Madigan (1968) indicated that both the low- and high- D_4 word lists had average imagery ratings of 5.2-5.7 on a seven-point scale where 7 indicates maximum concreteness. In other words, imagery apparently did not vary significantly in the lists of Experiment II.

Experiment III

The primary difference between the results of the present study on D_4 and those obtained by Rosen et al. (1962) on m was that the isolation effect was significant at both levels of D_4 , rather than only with low- D_4 materials. As a result, the expected interaction between isolation versus control lists and high- versus low- D_4 lists was not obtained.

Winograd's (1966) work, which demonstrated the apparently greater distinctiveness of high- D_4 items, had led to the expectation that the interaction would be significant. A possible explanation for the obtained results is contained in Kasschau and Pollio's comment (1967) on the

initial superiority of associative over semantic relations in mediating response transfer. Their suggestion was that subjects have at least two different types of word relations available, associative and semantic. Where only the semantic-similarity relation is available, words do not as rapidly benefit as they would from an associative-similarity relation; the subjects need more time to avail themselves of the semantic relationships. If this is so, the consistency of the significance of the isolation effect might have resulted from the lack of differentiation of the words composing the high- D_4 list, especially during the initial learning trials.

In conclusion, it can be said that with associative meaningfulness (m) held constant, variations in semantic meaningfulness (D_4) produced changes in performance consistent with the changes produced by similar variations in associative meaningfulness. The D_4 measure does, in fact, appear to measure an aspect of meaningfulness not detected by m , an aspect which influences performance even at relatively high levels of m .

Notes

The research reported was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. The assistance of Carol Ann Moore (Experiment I), John Shen (Experiment II), and James Kempson (Experiment III) in collecting and analyzing the data is gratefully acknowledged. Received for publication December 14, 1970.

References

- Anderson, N. H. 1963. Comparison of different populations: Resistance to extinction and transfer. *Psychological Review* 70:162-179.
- Bousfield, W. A. 1961. The problem of meaning in verbal learning. In *Verbal learning and verbal behavior*, ed. C. N. Cofer, pp. 81-91. New York: McGraw-Hill.
- Cieutat, V. J., Stockwell, F. E., and Noble, C. E. 1958. The interaction of ability and amount of practice with stimulus and response meaningfulness (m , m') in paired-associate learning. *Journal of Experimental Psychology* 56:193-202.
- Deese, J. 1965. *The structure of associations in language and thought*. Baltimore: The Johns Hopkins Press.
- Howe, E. S. 1965. Uncertainty and other associative correlates of Osgood's D_4 . *Journal of Verbal Learning and Verbal Behavior* 4:498-509.
- Kasschau, R. A. 1970. Unipolar vs. bipolar semantic-differential rating scales. Submitted to *Journal of Experimental Psychology*.
- Kasschau, R. A., and Pollio, H. R. 1967. Response transfer mediated by meaningfully similar and associated stimuli using a separate-lists design. *Journal of Experimental Psychology* 74:146-148.

- Koen, F. 1962. Polarization, m , and emotionality in words. *Journal of Verbal Learning and Verbal Behavior* 1:183-187.
- Noble, C. E. 1952. The role of stimulus meaning (m) in serial verbal learning. *Journal of Experimental Psychology* 43:437-446, 44:465.
- Osgood, C. E. 1961. Comments on Professor Bousfield's paper. In *Verbal learning and verbal behavior*, ed. C. N. Cofer, pp. 91-106. New York: McGraw-Hill.
- Osgood, C. E., Suci, G. J., and Tannenbaum, P. H. 1957. *The measurement of meaning*. Urbana: University of Illinois Press.
- Paivio, A. 1968. A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behavior* 7:41-49.
- Paivio, A., Yuille, J. C., and Madigan, S. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology* 76 (1, part 2).
- Rosen, H., Richardson, D. H., and Saltz, E. 1962. Supplementary report: Meaningfulness as a differentiation variable in the von Restorff effect. *Journal of Experimental Psychology* 64:327-328.
- Spreen, O., and Schulz, R. W. 1966. Parameters of abstraction, meaningfulness and pronunciability for 329 nouns. *Journal of Verbal Learning and Verbal Behavior* 5:459-468.
- Wallace, W. P. 1965. Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin* 63:410-424.
- Winograd, E. 1966. Recognition memory and recall as a function of degree of polarization on the semantic differential. *Journal of Verbal Learning and Verbal Behavior* 5:566-571.
- Zippel, B. 1967. Semantic differential measures of meaningfulness and agreement of meaning. *Journal of Verbal Learning and Verbal Behavior* 6:222-225.

continued from page 20

- Dethier, Vincent G. *Topics in Animal Behavior* by BIO. New York: Harper & Row, 1971. Pp. 77. Paperback, \$1.95.
- Dethier, Vincent G.; Lash, James W.; Lewontin, Richard C.; Lloyd, Monte; Iris, Hans; Roman, Hersche; Sussman, Alfred S.; and Telfer, William H. *Topics in the Study of Life: The BIO Source Book*. New York: Harper & Row, 1971. Pp. xiii, 482.
- Dockrell, W. B., ed. *The Toronto Symposium on Intelligence, 1969*. New York: Barnes & Noble, 1971. Pp. 267. \$9.50.
- Ehrenberg, A. S. C., and Pyatt, G. F., eds. *Consumer Behaviour*. Baltimore: Penguin Books, 1971. Pp. 384. Paperback, \$3.95.
- Elder, Glen J., Jr. *Adolescent Socialization and Personality Development*. Chicago: Rand McNally, 1971. Pp. 132. Paperback, \$2.50.
- Ellis, Norman R., ed. *International Review of Research in Mental Retardation*. Vol. 5. New York: Academic Press, 1971. Pp. xv, 408. \$17.00.
- Englehardt, Karlheinz. *Der Patient in seiner Krankheit*. Stuttgart: Georg Thieme, 1971. Pp. viii, 143.
- Epstein, William, and Shontz, Franklin. *Psychology in Progress*. New York: Holt, Rinehart and Winston, 1971. Pp. xi, 428.
- Esser, Aristide H., ed. *Behavior and Environment: The Use of Space by Animals and Men*. New York: Plenum Publishing, 1971. Pp. xvii, 411. \$17.50.
- Evan, William M., ed. *Organizational Experiments: Laboratory and Field Research*. New York: Harper & Row, 1971. Pp. ix, 274. \$8.00.
- Ewen, Robert B. *Introductory Statistics for the Behavioral Sciences*. New York: Academic Press, 1971. Pp. xiv, 155.
- Ewen, Robert B. *Introductory Statistics for the Behavioral Sciences Workbook*. New York: Academic Press, 1971. Pp. xiv, 155. Paperback, \$2.95.
- Eysenck, H. J., ed. *Readings in Extraversion-Introversion: Two Fields of Application*. London: Granada Publishing, 1971. Pp. 355. 99s.
- Ferguson, George A. *Statistical Analysis in Psychology and Education*. New York: McGraw-Hill, 1971. Pp. xii, 492. \$10.95.
- Fraisse, Paul, ed. *L'année psychologique*. Paris: Presses Universitaires de France, 1970. Pp. 345, 691. Fr42.
- Freedman, Jonathan L.; Carlsmith, J. Merrill; and Sears, David O., eds. *Readings in Social Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1971. Pp. xii, 573. Paperback, \$6.95.
- Friedman, Leonard J. *Virgin Wives: A Study of Unconsummated Marriages*. New York: Barnes & Noble, 1971. Pp. xiii, 161. Paperback, \$2.00.
- Galdston, Iago, ed. *The Interface between Psychiatry and Anthropology*. New York: Brunner/Mazel, 1971. Pp. xii, 150. \$7.50.
- Gersuni, G. V., ed. *Sensory Processes at the Neuronal and Behavioral Levels*. New York: Academic Press, 1971. Pp. ix, 333. \$18.00.
- Goguelin, Pierre; Cavozi, Jean; Dubost, Jean; and Enriques, Eugene. *La formation psychosociale dans les organisations*. Paris: Presses Universitaires de France, 1971. Pp. 204. Fr12.

continued on page 62

OVERT/COVERT PRONUNCIATION AND REHEARSAL IN VERBAL-DISCRIMINATION LEARNING

Lawrence E. Cole and N. Jack Kanak

University of Oklahoma

An experiment on the effects of overt versus covert rehearsal and pronunciation of the correct response. Neither variable influenced the rate or difficulty of verbal-discrimination learning, but overt pronunciation facilitated the free recall of 'correct' items whereas covert pronunciation plus overt rehearsal produced inferior free recall of 'incorrect' items.

The frequency theory of verbal-discrimination learning holds that subjects discriminate the right item from the wrong item of a pair on the basis of the differential subjective experience of the frequency of the two items (Ekstrand, Wallace, and Underwood, 1966). Ekstrand et al. assumed, as a working hypothesis, that qualitative differences do not exist in the sources of frequency units (that is, in *representational* responses, *pronunciational* responses, or *rehearsals* of the correct responses) with regard to their contribution to the differential-frequency cue. The present study explored the additional assumption, implicit in the literature, that differences do not exist in the saliency of a frequency unit as a function of whether the subject's mode of response for a pronunciation or a rehearsal response is covert or overt.

In the typical anticipation method, the subject indicates his selection during the anticipation interval by pronouncing aloud the item he has chosen. In the feedback interval, the pair reappears with the correct item underlined. Frequency theory assumes that the subject then covertly rehearses this 'right' item at least once during the feedback interval. Some studies (Ekstrand et al., 1966), however, have used a procedure whereby the overt pronunciation is eliminated; others (Kausler and Sardello, 1967) have required an overt rehearsal. Whether a covert pronunciation results in the same increment to the frequency cue as an overt one, and whether a covert rehearsal equals an overt one, has not been determined. The present study was directed toward these questions as well to the matter of whether overt/covert pronunciations produce the same effects as overt/covert rehearsals. A free-recall measure

was administered following acquisition to test whether availability of wrong and right items is affected by the mode of response for pronunciation and rehearsal responses.

In one condition (group O-C), the standard procedures used with the anticipation method were employed such that the pronunciation was overt and the rehearsal was assumed to occur at least once in a covert manner. A second condition (group C-C) required the subject to indicate his choice of the correct item by calling out 'left' or 'right.' It was assumed that in this procedure he would covertly pronounce the item prior to indicating its spatial position. A third condition (group O-O) involved overt pronunciation and overt rehearsal. Each subject here was asked to pronounce the correct item aloud once during the feedback interval. Finally, group C-O was produced by requiring the subject to make a spatial-position response in the anticipation interval and an overt rehearsal during the feedback interval.

METHOD

—Subjects—The 48 subjects were volunteers from introductory psychology classes at the University of Oklahoma; they received extra credit for participation. Each was randomly assigned to one of the four treatment conditions ($N = 12$) upon appearance at the laboratory, within the restriction that the number of subjects in any treatment condition at no time exceed that in any other condition by more than one.

—Materials—The list consisted of 15 pairs of associatively unrelated words selected from the Palermo-Jenkins (1964) norms. The right (correct) and wrong items and their specific pairings for list A were randomly determined. List B was a random re-pairing of list A within the restriction that no list A pairings be retained. Two additional lists (C and D) were created by reversing the item functions in lists A and B respectively. The four list variations were employed to control for differential ease in selecting the correct items, to minimize idiosyncratic associations, and to increase the generality of the findings. The use of the four lists was counterbalanced across subjects within each treatment condition. Four random orders of presentation and unsystematic variation of spatial position of wrong and right items within each serial random order were employed to minimize serial- and spatial-position cues.

—Procedure—The verbal-discrimination list was presented via a Lafayette memory drum at a 2:2-sec exposure rate with a 4-sec intertrial interval. Several examples of nonexperimental pairs were presented to ensure the subjects' comprehension of the instructions in each treatment condition.

The subjects were instructed to guess on the initial test trial. Practice was carried to one perfect trial, after which the free-recall task was administered. Here the subject was presented a sheet of paper with 30 lines (vertically

arranged, to discourage intrapair associative-recall strategies) and was instructed to write as many items as he could recall regardless of item function or order.

RESULTS AND DISCUSSION

Preliminary analyses of the numbers of trials to criterion and of total errors to criterion indicated that the four list variants were not significantly different in degree of difficulty (all $ps > .20$). The list factor was therefore omitted in all subsequent analyses.

The analyses of variance of trials to criterion (grand $\bar{X} = 7.19$) and of total errors to criterion (grand $\bar{X} = 24.80$) revealed no significant effects for either the mode of pronunciation or rehearsal or for the interactions (F typically less than 1.00). The means and standard deviations, respectively, for trials to criterion were 5.75 and 2.09 for group C-O, 7.42 and 3.47 for group O-O, 7.75 and 3.08 for group C-C, and 7.83 and 3.71 for group O-C. These results suggest that the mode of producing pronunciations or rehearsals has no differential effect on the saliency of the frequency cue.

In the free-recall task the means and standard deviations for wrong-item recall were 2.58 and 1.38 for group C-O, 5.58 and 2.97 for group C-C, 6.50 and 2.31 for group O-O, and 6.75 and 2.30 for group O-C, while the comparable values for right-item recall were 8.33 and 1.87 for group C-C, 8.67 and 2.06 for group C-O, 10.42 and 1.93 for group O-O, and 10.42 and 1.83 for group O-C.

A similar analysis of variance that included as a third factor the within-subject variable of type of item (wrong or right) recalled was performed to examine the possible influence of covert/overt pronunciations and rehearsals on free recall. The mode of the pronunciations frequency unit was significant [$F(1,44) = 21.18, p < .005$], with superior recall when the pronunciation was overt. The effect for mode of the rehearsal was nonsignificant ($p > .20$). The effect for type of item recalled was highly significant [$F(1,44) = 118.20$], but the interaction of mode of rehearsal and type of item was also significant [$F(1,44) = 5.63, p < .025$], as was the second-order interaction [$F(1,44) = 4.16, p < .05$]. Because of the second-order interaction, group differences in wrong- versus right-item recall were examined through separate analyses of variance for each type of item with the modes of the pronunciations and rehearsals as main factors.

For wrong-item recall both main effects were significant, with superior recall in each case for overt responding. However, in addition, the inter-

action of mode of pronunciation and mode of rehearsal was significant [$F(1,44) = 4.25, p < .05$]. In the statistical analysis of the interaction, group C-O was significantly inferior to each of the other conditions. Apparently a covert pronunciation combined with an overt rehearsal results in the subject's directing his attention during the feedback interval to the correct item at the expense of the wrong one. The 'left' or 'right' verbal response used to ensure a covert pronunciation may effectively prevent the subject from encoding the wrong item, thus making it more difficult to retrieve. Although group C-O required fewest trials to criterion ($\bar{X} = 5.75$), the overall effects of either mode variable on acquisition were nonsignificant.

In the comparable analysis of right-item recall, the mode of pronunciation was significant [$F(1,44) = 11.88, p < .005$], with overt pronunciations also resulting in substantially greater availability of right (correct) items. In contrast to the effects observed in wrong-item recall, neither the mode of rehearsal nor the interaction of modes of pronunciation and rehearsal was significant in right-item recall ($F_s < 1$). Thus the mode of pronunciation had a pronounced effect upon the availability of these correct items independent of the mode of rehearsal. Perhaps the manipulation designed to produce covert pronunciations suppressed the accrual of right-item availability. At least, (a) overt rehearsal failed to produce such facilitation and (b) the level of right-item free recall (approximately 50%) in the covert-pronunciation conditions was substantially below the 70%-75% level typically found in other studies using the usual covert-rehearsal procedure with similar materials (Fulkerson and Kausler, 1969).

While the present results suggest that covert and overt pronunciation responses and rehearsals of the correct responses have comparable effects on acquisition in verbal discrimination, the generality of this conclusion should be demonstrated with other methods of producing covert pronunciation. An improvement on the design of the present study might employ an identical manipulation for both covert pronunciations and rehearsals.

Notes

The authors thank Fred Landman for collecting the data, and John McCullers for his critical reading of an earlier version of the manuscript. Requests for reprints should be addressed to the first author, now at the Department of Psychology, Arkansas Polytechnic College, Russellville, Arkansas 72801. Received for publication January 4, 1971.

References

- Ekstrand, B. R., Wallace, W. P., and Underwood, B. J. 1966. A frequency theory of verbal-discrimination learning. *Psychological Review* 73:566-578.
- Fulkerson, F. E., and Kausler, D. H. 1969. Effects of intrapair and interpair bidirectional associates on verbal-discrimination learning. *Journal of Verbal Learning and Verbal Behavior* 8:307-310.
- Kausler, D. H., and Sardello, R. J. 1967. Item recall in verbal-discrimination learning as related to pronunciation and degree of practice. *Psychonomic Science* 7:285-286.
- Palermo, D. S., and Jenkins, J. J. 1964. *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press.

continued from page 56

- Grad, Frank P.; Goldberg, Audrey L.; and Shapiro, Barbara A. *Alcoholism and the Law*. Dobbs Ferry, N.Y.: Oceana Publications, 1971. Pp. xviii, 311. \$15.00.
- Hall, John F. *Verbal Learning and Retention*. Philadelphia: J. B. Lippincott, 1971. Pp. xiv, 590. \$12.95.
- Hamblin, Rolbert L.; Buckholdt, David; Ferritor, Daniel; Kozloff, Martin; and Blackwell, Lois. *The Humanization Processes: A Social Behavioral Analysis of Children's Problems*. New York: Wiley, 1971. Pp. xxi, 305. \$12.50.
- Hellmuth, Jerome, ed. *Cognitive Studies*. Vol. 2, *Deficits in Cognition*. New York: Brunner/Mazel, 1971. Pp. x, 503. \$15.00.
- Hellmuth, Jerome, ed. *Exceptional Infant*. Vol. 2, *Studies in Abnormalities*. New York: Brunner/Mazel, 1971. Pp. ix, 529. \$15.00.
- Herbst, P. G. *Behavioural Worlds: The Study of Single Cases*. New York: Barnes & Noble, 1971. Pp. xiv, 248. \$8.00.
- Herriot, Peter. *Language and Teaching: A Psychological View*. New York: Barnes & Noble, 1971. Pp. vii, 109. \$4.00; paperback, \$2.00.
- Horton, David L., and Jenkins, James J., eds. *The Perception of Language*. Columbus, Ohio: Charles E. Merrill, 1971. Pp. xix, 267. \$8.95.
- Howard, James A. *The Ego Mill: Five Cases in Clinical Psychology*. Chicago: Cowles, 1971. Pp. x, 245. \$6.95.
- Howells, John G., ed. *Modern Perspectives in World Psychiatry*. New York: Brunner/Mazel, 1971. Pp. xxxii, 787. \$25.00.
- Irwin, Francis I. *Intentional Behavior and Motivation: A Cognitive Theory*. Philadelphia: J. B. Lippincott, 1971. Pp. xv, 201. \$8.95.
- Kaneff, S., ed. *Picture Language Machines: Proceedings of a Conference held at the Australian National University, Canberra, on February 24-28, 1969*. London: Academic Press, 1970. Pp. xiv, 425. £4 50s.
- Kantor, J. R. *The Aim and Progress of Psychology and Other Sciences*. Granville, Ohio: Principia Press, 1971. Pp. ix, 613.
- Keezer, William S. *Mental Health and Human Behavior*. 3d ed. Dubuque, Iowa: Wm. C. Brown, 1971. Pp. xiv, 298. Paperback, \$4.95.
- Kipnis, David. *Character Structure and Impulsiveness*. New York: Academic Press, 1971. Pp. xi, 133. \$7.95.
- Klein, Melanie; Heimann, Paula; and Money-Kyrle, Roger, eds. *New Directions in Psycho-analysis: The Significance of Infant Conflict in the Pattern of Adult Behaviour*. New York: Barnes & Noble, 1971. Pp. xiii, 534. Paperback, \$3.50.
- Lambert, William Wilson, and Weisbrod, Rita. *Comparative Perspectives on Social Psychology*. Boston: Little, Brown and Co., 1971. Pp. 309. Paperback, \$4.95.
- Lasswell, Harold D. *Propaganda Technique in World War I*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xxxii, 233. Paperback, \$2.95.
- LeCron, Leslie M. *The Complete Guide to Hypnosis*. Los Angeles: Nash Publishing, 1971. Pp. 225. \$6.95.
- Lehrman, Daniel S.; Hinde, Robert A.; and Shaw, Evelyn, eds. *Advances in the Study of Behavior*. Vol. 3. New York: Academic Press, 1971. Pp. xi, 263. \$12.00.

continued on page 80

SUBVOCALIZATION OF HEARD OR SEEN WORDS PRIOR TO SPOKEN OR WRITTEN RECALL

John L. Locke and Fred S. Fehr

Children's Research Center, Champaign, Illinois

As 24 adults tried to recall words lists presented visually ('see') or aurally ('hear') with oral ('say') or written ('hear') responses, EMG recordings measured covert phonetic activity on recorded labial and nonlabial phonemes. During stimulus presentation, subvocalization occurred only in conditions 'see-say' and 'hear-write,' which required translating the stimuli into a different mode and code for response. During recall rehearsal, it occurred in all four conditions.

The role of so-called subvocal speech in thinking, learning, and reading has been the subject of numerous studies in the last century (see reviews in Edfeldt, 1960; Locke, 1970; McGuigan, 1970). Surprisingly, very few of these investigations asked whether subvocal speech actually is a covert form of overt (communicative) speech, and none were able to establish that it is (Locke and Fehr, 1970a).

Recently, however, a series of experiments (Locke and Fehr, 1970a, 1970b; Goldstein, Locke, and Fehr, 1972; Locke, 1971) presented subjects with phonemically classed materials while recording the electromyographic (EMG) potentials arising from vocal-tract sites sensitive to the covert processing of these classes. Specifically, the visual stimuli had labels or pronunciations that did or did not require the articulation of labial phonemes [p, b, m, w, f, v]; surface electrodes at the chin and lower lip reflected the relative activity of labial musculature. The general finding was that when subjects were rehearsing labial items, their EMG tracings were greater in amplitude than when they were processing nonlabial items. As a result of this work, it now appears that subvocalization is a covert form of speech, occurs in children and adults, is elicited by orthographic and pictorial stimuli, and accompanies both the prerecall processing of isolated words and the silent reading of connected language.¹ However, several questions that arose during the conduct of these studies remain to be answered.

The degree of children's phonetic coding activity during item presentation has been observed to be nearly identical for pictures and printed

words (Goldstein et al., 1972). Both forms of nonphonetic stimuli were converted into phonetic terms in their prerecall processing. Were the stimuli covertly articulated for the purpose of recoding them into phonetic features, which might be more resistant to decay or interference than nonphonetic cues? Or were they subvocally rehearsed primarily to maintain them in immediate memory, with speech serving simply as an accessible repertoire of repeatable motor-acoustic units? Presumably, such a distinction could be made by presenting subjects with phonetic and nonphonetic stimuli, intramodal if possible.

Covert phonetic activity is not quite as vigorous during the presentation of visual stimuli as it is during a subsequent delay, or rehearsal, period preceding oral recall (Locke and Fehr, 1970b). Is this increase in phonetic processing attributable mostly to the disappearance of the stimuli and the subject's attendant struggle to keep them in immediate memory? Or is at least some of this subvocalization an anticipation and practice of the oral response itself?

This paper reports an experiment designed to identify the more likely reasons for phonetic coding activity. Essentially, it involved the analysis of EMG-recorded covert phonetic behavior during the presentation of phonetic and nonphonetic stimuli and immediately prior to their orally or graphically registered recall.

METHOD

—Subjects—The 24 adults, between the ages of 15 and 51, served on a voluntary basis. All were naive to the experimental purpose. Educational background varied from partial completion of high school to completion of a master's degree.

—Stimuli—The stimuli were eight lists of five words whose pronunciation requires the articulation of labial phonemes and eight lists of five words whose pronunciation contains no phonemes requiring labial articulation. Items were matched across lists for syllable length and written frequency (Thorndike and Lorge, 1944). Most of the words were monosyllabic and concrete, and all but 15% were A or AA according to the frequency count. Items appeared on 35-mm slides in uppercase type. Two labial lists and two nonlabial lists were randomly assigned to each of four subsets of 20 words.

—Procedure—The subjects were told they were in an experiment on the physiology of learning, and they were seated individually approximately 10 ft from a slide screen and 3 ft from a loudspeaker (Sony TC-124). A response sheet was placed on a lapboard, and the microphone of an intercom system was situated directly behind each subject's head. The subjects were instructed to observe and recall, in sequence, lists of familiar words. They were also

told of the two stimulus-presentation conditions and the two response modes, so that they knew whether they would see words and orally ('see-say') or graphically ('see-write') recall, or would hear words and orally ('hear-say') or graphically ('hear-write') respond. Each subject was in all four conditions.

Each experimental condition contained two labial (L) lists, two nonlabial (N) lists, and two control (C) lists in which the subject either saw five flashing lights (blank slides) or heard five sounds (500-Hz pulses), both of which corresponded in distribution and duration to the verbal stimuli. Half the subjects got the lists in a N L C N L C order; the other 12 subjects had a L N C L N C sequence.

The exposure time of aural stimuli, presented at a comfortable listening level via cassette tape recorder and speaker (Sony TC-124), ranged from 180 to 680 msec ($\bar{X} = 418$) with interstimulus intervals of 600 to 1,120 msec ($\bar{X} = 864$). Visual stimuli, presented by an automated slide projector (Kodak Carousel AV-900), ranged from 400 to 480 msec ($\bar{X} = 456$) with 780- to 820-msec interstimulus intervals ($\bar{X} = 803$).

Each series of five stimuli (presentation period) was followed by a 10-sec delay period (rehearsal period) in which the subjects were to await the opportunity to recall. In both 'see' conditions the signal to respond was a red light for 10 sec. In the 'hear' conditions, the subjects were cued by a 125-Hz continuous tone of the same delayed onset and duration. They were to recall only when the light or tone was on. The recall period was followed by a 20-sec period in which the subjects simply awaited the next presentation period, an interval included to permit a reduction in residual oral tension after oral recall. In the control condition, the subjects were to write or say "I heard five beeping sounds" or "I saw five flashing lights" while the signal tone or light was on. Primarily, however, the purpose of the control condition was to test a hypothesis associated with an unrelated co-occurring study; indeed, previous work has shown control values to be less than labial and nonlabial tracings (Locke and Fehr, 1970a) and has suggested that it is appropriate to consider the *nonlabial* condition as a control above which labial tracings must rise if oral activity is phonetic in nature. Lastly, one of the 24 possible sequences of the four conditions was assigned to a subject, and the four subsets of 20 words (4 lists $\times$ 5 items) were used in each condition six times, removing the possible problem of variable list difficulty.

Before testing, two surface electrodes were attached to the chin and inferior surface of the lower lip of each subject. This placement has been found in previous work (Locke and Fehr, 1970a, 1970b; Goldstein et al., 1972; Locke, 1971) to be sensitive to surface tension in the region of the lower lip and to distinguish labial from nonlabial activity, thereby indicating the presence of subvocalization on the labial phonemes in the present design. Additional leads were placed on the left ankle, two fingers of the left hand, and below the right temple as part of a separate experiment and to reduce the subject's attention to the chin-lip placement and any subsequent deduction that the experiment was related to speech. The Grass 7 polygraph (7P3A preamplifier and integrator and filters) that recorded EMG activity, the slide projector, and the tape recorder were all in a control booth. The subjects were instructed over the intercom at the completion of each experimental condition. Test time was approximately 30 min.

RESULTS

Analysis was begun by identifying the single greatest tracing peak within each subject's presentation and rehearsal periods for labial, nonlabial, and control stimuli in all lists of the four conditions. These were statistically analyzed to determine whether the labial peaks were significantly greater than the nonlabial ones by a priori t tests (after Winer, 1962, p. 85). In all cases, both labial and nonlabial presentation-period values also significantly ($p < .05$) exceeded control peaks. Table 1 shows that labial-nonlabial discrepancies during the presentation period were significant for the 'see-say' and 'hear-write' conditions, and for all four conditions during the rehearsal period—subvocalization did occur, as the t values show.

Table 1. Mean EMG peak values in microvolts for labial and nonlabial tracings taken during presentation and rehearsal of words in the 'see-say' (S-S), 'see-write' (S-W), 'hear-say' (H-S), and 'hear-write' (H-W) conditions

	Presentation				Rehearsal			
	S-S	S-W	H-S	H-W	S-S	S-W	H-S	H-W
Labial	54.9	47.0	40.4	40.1	55.2	46.4	49.3	42.2
Nonlabial	43.2	42.7	41.9	35.5	42.6	36.8	41.3	34.8
r	.66‡	.66‡	.73‡	.88‡	.69‡	.87‡	.79‡	.84‡
t	2.86‡	1.10	.55	2.10*	3.40‡	3.73‡	2.63‡	2.60‡

* $p < .05$. † $p < .01$. ‡ $p < .001$.

In order to preserve the labial-nonlabial relationships in these data, between-condition analyses used labial-to-nonlabial, or L/N, ratios (as in Goldstein et al., 1972), where entries greater than 1.0 are considered suggestive of phonetic activity. A one-way analysis of variance was computed on these ratios for presentation and rehearsal periods. A significant F was found for presentation [$F(3, 92) = 3.16$; $p < .01$] but not for rehearsal [$F(3, 92) = 0.56$; $p > .05$] periods. Subsequent tests of simple effects among presentation data (as described by Ferguson, 1966, p. 296) showed that the 'see-say' ratio mean ($L/N = 1.44$) significantly exceeded that for 'see-write' [$L/N = 1.13$; $t = 2.21$; $p < .05$] and 'hear-say' [$L/N = 1.00$; $t = 3.08$; $p < .01$] but not 'hear-write' [$L/N = 1.20$; $t = 1.65$; $p > .05$] conditions. The superiority of these 'hear-write' over 'hear-say' values approached significance [$t = 1.39$; $p < .20$]. Analyses of differences between the presentation and rehearsal data were not conducted because visual inspection showed the ratio discrepancies to be of negligible magnitude ('see-say,' $-.11$; 'see-write,' $.14$; 'hear-say,'

.18; 'hear-write,' .00). Recall scores for all conditions exceeded a mean of 4.5/5 items in a list and did not differ significantly.

DISCUSSION

The data, particularly those on stimulus presentation, lend themselves to several interpretations. In stimulus presentation, the only substantial evidence of phonetic coding activity was in conditions where processing was required to cross a mode and code (in condition 'see-say,' with visual-oral modes and orthographic-phonetic codes, and in condition 'hear-write,' with aural-visual modes and phonetic-orthographic codes). In conditions where no crossing of mode and code was required (in condition 'see-write,' with visual-visual modes and orthographic-orthographic codes, and in condition 'hear-say,' with aural-oral modes and phonetic-phonetic codes), there was no evidence of phonetic activity during stimulus presentation.

Since rehearsal in all conditions was probably phonetic in nature (Locke and Fehr, 1970a), the greater amount of phonetic coding in the 'see-say' than in the 'see-write' condition during stimulus presentation may have been due to the subjects' knowledge of the mode in which the responses were to be made. The greater phonetic activity in the 'see-say' than in the 'hear-write' condition (see the *t* values in Table 1) may also have been attributable to such a preparatory recruitment of vocal-tract musculature, but there also may well have been some covert phonetic activity associated with the recoding of orthographic stimulus sequences. That is, even when subjects are engaged in silent reading, they show greater labial than nonlabial tracings (Locke, 1971), a showing thought by some investigators to be a residual artifact of earlier phonics training (Goldstein et al., 1972) in which printed words are recoded into their meaningful phonological correlates. The tendency of 'hear-write' values to exceed those obtained under the 'hear-say' conditions seems to underscore the apparent significance of stimulus recoding. It might also be observed that the fact that neither 'say' condition increased significantly as the response period approached does not provide support for the anticipation hypothesis.

Without visual-phonetic stimulus and response modes—that is, without presenting phonetically transcribed material and without requiring transcription into phonetic symbols in responding—it was not possible to achieve symmetry in the experimental design and keep mode and code separated. If the stimulus was visual, it was nonphonetic; if a non-oral response was required, it was of necessity nonphonetic too. But

phonetic symbols were purposely avoided, since they probably would have elicited unusually high levels of subvocal activity (even in phoneticians), a common finding when subjects are presented with a relatively unfamiliar code (Edfeldt, 1960).

In sum, the most interesting finding was that subjects subvocalized more in the presentation periods of the 'see-say' and 'hear-write' than 'see-write' and 'hear-say' conditions, which seems to suggest stimulus recoding during stimulus exposure. It also suggests that the mode and code of the stimulus and the response are not nearly as evocative of covert phonetic activity as is the *translation* of mode and code from stimulus to response.

Notes

This paper was supported by Public Health Research Grant MH07346 from the National Institute of Mental Health. Revision received for publication April 28, 1971.

1. An analogous form of behavior is displayed by deaf children who rehearse orthographic stimuli dactylally (using fingerspelling gestures) and show dactylic mediation effects in recall from short-term memory (Locke and Locke, 1971).

References

- Edfeldt, A. W. 1960. *Silent speech and silent reading*. Chicago: University of Chicago Press.
- Ferguson, G. A. 1966. *Statistical analysis in psychology and education*. New York: McGraw-Hill.
- Goldstein, J., Locke, J. L., and Fehr, F. S. 1972. Children's pre-recall phonetic processing of pictures and printed words. *Psychonomic Science*, in press.
- Locke, J. L. 1970. Subvocal speech and speech. *Asha* 12:7-14.
- Locke, J. L. 1971. Phonemic processing in silent reading. *Perceptual and Motor Skills* 32:905-906.
- Locke, J. L., and Fehr, F. S. 1970a. Subvocal rehearsal as a form of speech. *Journal of Verbal Learning and Verbal Behavior* 9:495-498.
- Locke, J. L., and Fehr, F. S. 1970b. Young children's use of the speech code in a recall task. *Journal of Experimental Child Psychology* 10:367-373.
- Locke, J. L., and Locke, V. L. 1971. Deaf children's phonetic, visual, and dactylic coding in a grapheme recall task. *Journal of Experimental Psychology* 89:142-146.
- McGuigan, F. J. 1970. Covert oral behavior during the silent performance of language tasks. *Psychological Bulletin* 74:309-326.
- Thorndike, E. L., and Lorge, I. 1944. *The teacher's word book of 30,000 words*. New York: Columbia University Teacher's College Bureau of Publications.
- Winer, B. J. 1962. *Statistical principles in experimental design*. New York: McGraw-Hill.

TOTAL TIME IN VERBAL-DISCRIMINATION LEARNING

John H. Mueller and John L. Flanagan

University of Missouri, Columbia

Anticipation and feedback intervals were factorially lengthened or shortened over trials so that all groups had the same total learning time. The performance of those groups at the end of that time was equivalent, thus generally confirming the total-time hypothesis. Displaced rehearsal and learning to learn are discussed as possible factors in the results. Item recall and associative matching seemed most affected by temporal variation early in learning.

The total-time hypothesis is that the total amount of time required to master a set of materials remains invariant across differences in rate of presentation—that although the number of trials to criterion may vary between groups with different presentation rates, the time per item multiplied by the number of trials generally produces a comparable figure in all groups. A recent review suggested that the hypothesis holds provided the task involves no more than rehearsal (Cooper and Pantle, 1967). It also noted that variation of the stimulus duration (anticipation in paired associates) frequently contradicts the hypothesis, whereas variation of the feedback interval is less likely to produce a difference in total time.

The present experiment was a test of the total-time hypothesis in verbal-discrimination learning. This learning situation involves presenting a series of word pairs to the subject and requiring him to learn which member of each pair has been designated as right (R-term). A current explanation assumes that the recognition of R-terms involves a difference in the subjective frequency of experience for R-terms relative to the wrong members (W-terms) in each pair (Ekstrand, Wallace, and Underwood, 1966). The theory assumes that the subject's rehearsal of R-terms during feedback eventually provides him with a functional cue to distinguish R-terms from W-terms. Since the R-term has been more frequently experienced due to that rehearsal, the subject need only choose the most frequently experienced word in each pair to respond correctly.

Cooper and Pantle (1967) did not consider the total-time hypothesis

in verbal-discrimination learning as such (only in paired-associate, free-recall, and serial learning) but it seems clear that the frequency-theory view would satisfy their boundary condition of 'rehearsal only' quite well, so that the hypothesis should hold in verbal-discrimination learning. The frequency-theory emphasis on the subject's rehearsal of the R-terms during the feedback interval also implies that longer *anticipation* intervals might produce no benefit in learning (in terms of trials to criterion) but just take more time, thereby contradicting the total-time hypothesis. Although the processes may differ, if the anticipation period in verbal-discrimination learning is seen as analogous to the stimulus duration in paired-associate learning, such a departure from the hypothesis would at least be consistent with the other tasks reviewed by Cooper and Pantle (1967). However, there should be no discrepancy from the total-time hypothesis due to *feedback* variation, according to the frequency-theory prediction, since more time here should lead to fewer trials, as the effects of extra rehearsal during the extended feedback durations offset those longer intervals.

The literature on these questions is ambiguous. Some studies have varied either the anticipation or feedback intervals (Ingison and Ekstrand, 1970; Radtke, McHewitt, and Jacoby, 1970; Skeen, 1970), and others have varied both simultaneously (Kanak, 1968), but a complete factorial seems desirable (Nodine and Goss, 1969). Farzanegan (1969) and Wike and Wike (1970) did use complete factorials: the latter study reported no total-time results; the former, that longer anticipation and feedback intervals both took more time, with no interaction. Similar exceptions to the total-time hypothesis have been found in many of the studies that varied only one or the other of the intervals. Since verbal-discrimination learning should optimally satisfy the rehearsal-only requirement, these exceptions indicate the need for further analysis.

It has been noted elsewhere (Mueller and Slaymaker, 1970) that the frequency-theory prediction of the effect of anticipation duration may have been oversimplified in the past. While it is true that frequency theory places the critical events in the feedback frame (where rehearsal initiates the frequency shift to R-terms), the eventual nonrandomness of choices should also be taken into account; that is, with rehearsal becoming possible during the *anticipation* period as the subject acquires the discrimination, the importance of the duration of the anticipation interval may change over the course of learning. Such displaced covert rehearsal has been noted in paired-associate learning (Zacks, 1969), and to the extent that it also occurs in verbal-discrimination learning, especially during the later anticipation periods, may be involved in the discrepancies from the total-time hypothesis.

Changing the duration of anticipation and feedback intervals during the course of learning (Mueller and Slaymaker, 1970) seemed the appropriate methodology. Briefly, for the same number of trials some subjects receive very short periods for anticipation early in learning, the periods lengthening over trials; other subjects receive long initial anticipation intervals, the intervals shortening over trials. The same manipulations are made for feedback duration within each of the two anticipation groups. If the terminal duration in the increasing groups is used as the starting duration in the decreasing groups, all groups have the same amount of time overall, though distributed differently. This method offers a test of the total-time hypothesis and of the notions (outlined above) about the importance of the anticipation interval in verbal-discrimination learning.

METHOD

—Design and subjects—The basic design was a 2×2 factorial for anticipation-interval duration (increasing, decreasing) and feedback-interval duration (increasing, decreasing), with one control group outside the factorial. Two different list forms were used, the R-terms in one being the W-terms in the other, and vice versa; these forms were used equally in all groups. The first 80 subjects served in partial fulfillment of course requirements in introductory psychology and were assigned to the factorial groups in rotation with one restriction: since sex has been found to be a factor in verbal-discrimination learning (Fulkerson and Johnson, 1971), men and women were equally distributed in all groups. The 20 subjects for the control group were then run, equally distributed over the list forms by sex.

—Materials and apparatus—A 20-pair list was used (the control list of Ekstrand et al., 1966, p. 572). The pairs were arranged in four different orders of presentation, each of which was used about equally often as the starting order in each group. Each word in all pairs appeared twice on the top and bottom during the anticipation frame across the orders, and the R-term was on the top and bottom for half of the pairs within each order. The two forms of the list were presented with a Stowe memory drum by the anticipation method.

—Procedure—With a few exceptions, the procedure was that of Mueller and Slaymaker (1970). All subjects were given eight trials on the list, with anticipation and feedback intervals as follows. Group II had both anticipation- and feedback-interval durations increasing (lengthening) over trials, group ID had the anticipation interval increasing and the feedback interval decreasing (shortening), group DI had anticipation decreasing and feedback increasing, and group DD had both intervals decreasing over trials. The increases or decreases were made every other trial, so that four durations were used: 1 sec, 2 sec, 3 sec, and 4 sec, with intertrial intervals of about 10 sec. For example, the ID group had a 1:4-sec rate for anticipation and feedback on both trials

1 and 2, a 2:3-sec rate on trials 3 and 4, and so on. The control group (CC) received constant anticipation and feedback durations for all training trials, with 2.5 sec for anticipation and feedback. Thus each group had the same amount of time at the end of eight trials, although distributed quite differently. The subjects were familiarized with the duration of the intervals by the experimenter's demonstrating the shutter operation on a blank part of the tape.

After the last trial, the subjects were given a final test trial, with as much time as they desired, to choose the R-term in each pair, but they got no feedback. They were then given two more unpaced tests, one of free recall and the other of associative matching. A sheet with 40 lines on it was given to each subject and he was told to write down all of the words he could remember in any order. No feedback was given. When the free-recall task was completed, the subject was given a sheet of paper with the R-terms in one column and the W-terms in another column, and asked to form the appropriate pairs.

RESULTS

Total time

Table 1 shows the means for several measures of acquisition: (a) the total number correct over the eight trials, (b) the number correct on the final, unpaced trial, (c) the number of pairs learned to a criterion of three consecutive errorless trials, and (d) the time used (in seconds) up to the criterion run. These measures of performance were analyzed using procedures described by Winer (1962) for comparing outside control groups to the factorial groups. Except where otherwise noted, all *F*s have 1 and 95 *df*.

The measure of *time used* was the only one of the acquisition measures to show a significant difference for either main effect, with long initial

Table 1. Summary statistics by group

Measure	Group				
	II	ID	DI	DD	CC
Total correct	129.0	132.0	129.9	133.5	138.5
Unpaced correct	19.8	19.7	19.8	19.5	19.8
Criterion pairs	18.3	19.0	18.3	18.8	19.3
Time used	139.7	178.8	171.5	206.0	130.4
R-term recall	12.1	12.4	13.1	14.1	14.6
W-term recall	8.7	9.9	10.4	10.7	11.3
Matching	11.4	14.9	15.5	15.3	14.3

Note: I, D, and C refer to whether the anticipation or feedback frames were increasing, decreasing, or constant in duration. Time used is given in seconds.

anticipation and feedback intervals (decreasing) producing the greatest amount of time [$F_s = 3.37$ and 5.23 , $ps < .10$ and $.05$], as if the longer intervals at the outset were not being maximally utilized. The interaction of anticipation and feedback duration was nonsignificant for all acquisition measures ($F_s < 1$). In general, however, since the measures of *total correct* and of *unpaced correct* are the most valid measures for the test of the total-time hypothesis (as the others are based on points at which time has not yet equalized), the hypothesis seems confirmed.

Considering sex and list form as additional factors in the analysis for these measures, however, reveals some interactions with anticipation and feedback duration that may qualify the primary data but which must be cautiously interpreted due to the smaller cells ($N = 5$). Women were better than men in terms of the number of their *criterion pairs* and of their unpaced correct responses [$F_s = 6.69$ and 3.78 , $ps < .05$ and $.10$], but there was no difference either for total number correct overall or for time used ($F_s < 1.83$). The interaction of anticipation duration and sex for the measures of total correct and time used, however, indicates that women performed better than men with increasing anticipation durations (short intervals early), while men did somewhat better with the decreasing anticipation durations (long intervals early) [$F_s = 5.69$ and 5.88 , $ps < .05$]. Other interactions with sex did not reach significance.

One form of the list was harder in terms of the total correct and time used [$F_s = 6.10$ and 6.15 , $ps < .05$]. List form also interacted with the type of feedback arrangement, with the more difficult list acquired slower under the increasing arrangement but not under the decreasing arrangement [$F_s = 6.85$, 5.39 , and 4.48 for total correct, criterion pairs, and time used, $ps < .05$]. This may qualify acceptance of the total-time hypothesis, suggesting that the phenomenon is in some way a function of the difficulty of the materials, or it may indicate sampling error. However, much of this effect involved the difficulty of the one form of the list for the first block of two trials, which could indicate a guessing bias on the very first trial. Since Fulkerson and Johnson (1971) found that guessing on the first trial was inferior to observing with anticipation, we tentatively attribute the present effects of that form of the list to a guessing bias.

Finally, a comparison of the measures of acquisition for the control group (CC) with those for all factorial groups revealed a significant difference for total correct and time used [$F_s = 7.92$ and 5.95 , $p < .05$], with the control group better than the mean of the factorial groups. Comparing that control group to the individual means for total correct indicated that group CC differed from both groups II and DI [$ts(95) >$

2.45, $ps < .05$]. For time used, group CC differed significantly only from group DD [$t(95) = 3.21$, $p < .01$]. Group CC was quite comparable to the mean of the factorial groups and the individual groups in terms of unpaced correct and criterion pairs.

Verbal discrimination

The test of the total-time hypothesis required data on overall performance, but a test of the possibility of displaced rehearsal requires data on the time course of the groups' acquisition. Figure 1 shows the mean number correct by two-trial blocks for all five groups. Adding blocks

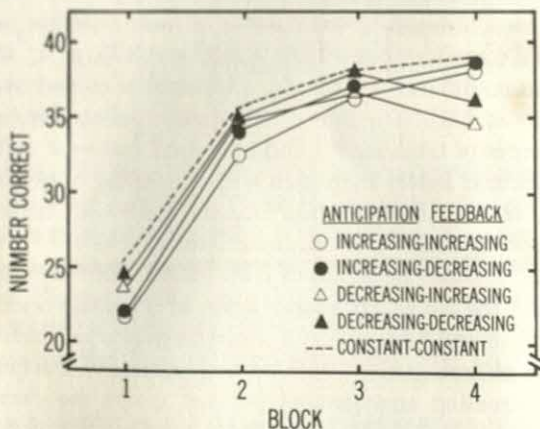


Fig. 1. Average number of correct responses by blocks of two trials for each group

to the factorial design revealed an interaction of anticipation duration and blocks, as groups with longer initial anticipation periods performed better than the groups with short initial anticipation periods (increasing) [$F(3, 192) = 20.76$, $p < .001$], especially early in learning. Although the reversal in the last block is consistent with the possibility of rehearsal in anticipation, since the 'increasing' groups had the longer intervals at that point, the relative performance of group CC indicates a performance handicap as the more likely interpretation. In any event, it is clear that long feedback intervals early in training did not significantly facilitate learning with either the long or short anticipation intervals. In fact, both groups ID and DD had longer feedback periods than group CC during blocks 1 and 2, yet failed to perform as well.

The longer anticipation intervals at the outset allowed the subjects to scan both alternatives more times, with the result that the frequency unit added to the R-term during feedback by rehearsal should not produce as large a difference with longer intervals: in effect, groups DD and DI should have had a 3:2 differential after the first trial, groups II and ID a 2:1 differential (see Ekstrand et al., 1966). That prediction is not supported in the data, however, as the longer anticipation intervals (2.5 and 4 sec) at the outset led to better performance than shorter ones (1 sec). It had been expected that longer anticipation intervals late in learning would allow displaced rehearsal, but the superiority of longer anticipation intervals seemed more general (within the factorial).

This advantage for longer anticipation periods throughout may be an artifact of the anticipation procedure, since the subjects could use the immediately following long anticipation period to offset a short feedback frame under the displaced-rehearsal interpretation. However, this advantage may also have been due, at least in part, to a performance handicap in the short anticipation-duration groups; if so, the handicap might be reduced somewhat with learning-to-learn experience. The slightly superior performance of group CC may reflect the reduction of this performance handicap, or it may be the result of the constancy of intervals allowing the subject to covertly displace his rehearsal more readily (Nodine, 1969).

Incidental learning

Table 1 also contains two supplementary measures taken after the training trials. The first measure showed that the *free recall* of both R- and W-terms was better when the early anticipation intervals were long (decreasing) [$F_s = 4.68$ and 3.86 , $ps < .05$ and $.10$]. That is, long anticipation intervals later did not compensate for short ones early in learning. The absence of a feedback effect ($F_s < 1.40$) suggests that the processes involved in the incidental ability to recall pair members depend primarily on the early anticipation periods—perhaps on the repeated scanning of alternatives before responding. Women again performed better than men [$F_s = 5.51$ and 3.45 , $ps < .05$ and $.10$, for R- and W-terms respectively], but with no interactions.

Another presentation of the results for free recall and a somewhat more detailed analysis of this free-recall data is shown in Table 2. Frequency theory implies that items that are correctly anticipated receive additional frequency units relative to items that are not correctly anticipated—receive those units because pronouncing and rehearsal responses both

occur. R-terms should generally be more available for recall than W-terms, as was found (Table 1), and it might also be expected that the R-term members of pairs that were frequently correct would be more available for recall than the R-term members of pairs that were less often anticipated correctly. These pairs will typically differ across subjects, of course, but the expectation of a gradient does not depend on commonality.

Since the trends were comparable, the groups were pooled, and the results for all 100 subjects are shown in Table 2. Pairs were categorized as a function of the number of times the pair was correct during acquisition. The probability of recall was computed by dividing the number of times the R- or W-term was in fact recalled by the number of pairs in that particular category (possible recalls).

Table 2. Probability of recall of R- and W-terms as a function of the number of times a pair was correct during acquisition

Times correct	R-term	W-term
2	.64	.60
3	.70	.65
4	.65	.61
5	.65	.50
6	.66	.54
7	.66	.51
8	.69	.45

Starting with pairs that were correct at least twice (to get adequate representation in the category), no systematic trend appeared for the R-terms, but the W-terms seemed somewhat less likely to be recalled the more times the pair was correct. If this indicates that the subject spends increasingly more time attending to R-terms when he is correct, then this decline for W-terms is consistent with frequency theory. The absence of an increase for R-terms is puzzling, but perhaps this would have appeared with further practice, though most subjects were approaching list criterion after eight trials.

The last row in Table 1 concerns another incidental measure, *associative matching*, taken to assess the extent to which the contiguity between R- and W-terms produced an incidental association. There was a reliable advantage for longer anticipation intervals early (decreasing) and, to a lesser extent, for long feedback intervals early as well [$F_s = 5.22$ and 2.67 , $ps < .05$ and $.15$]. The interaction of anticipation and feedback

was marginally significant [$F = 3.36, p < .10$], as the feedback effect was limited to the short initial anticipation periods. As with the other incidental measure (recall), it appears that long anticipation intervals occurring late in training did not encourage the development of association as did long anticipation intervals early in training. Since group II generally compared favorably to the other groups, these results seemingly cannot be attributed to differences in the degree of learning. Therefore, the recall and matching results both seem to indicate that incidental processes depend primarily on the early stages of verbal-discrimination learning.

DISCUSSION

The present findings for verbal-discrimination learning seem largely in accord with the total-time hypothesis. The two basic measures, total correct and the number correct on the unpaced trial, failed to reveal any simple effects due to anticipation or feedback variation, although there were interactions, with sex and list form, that are difficult to interpret. The superior acquisition of a control group given constant time intervals throughout seems best attributed to a performance factor, especially since this group performed no better than the others on the unpaced test, the latter presumably eliminating any performance handicap, but it may also have been easier for the subjects in this group to displace their rehearsal (Nodine, 1969). This slight superiority was also found in a previous study (Mueller and Slaymaker, 1970). Thus the total-time hypothesis seems confirmed, since all experimental groups performed comparably at the end (to indicate that the same amount had been learned).

There is a consideration that might somewhat qualify this conclusion. That is, since many subjects were approaching asymptotic performance, time used may somewhat reflect time to learning, as in trials-to-criterion tests. That seems rather tenuous, however, since not all subjects had reached list criterion, so a tentative acceptance of the hypothesis as confirmed seems more warranted.

As measures of acquisition performance in general, criterion pairs and time used seem quite acceptable, and should provide information about the time course of learning. Here, time used showed differences for both anticipation and feedback durations, with short early intervals eventually leading to less time, as if the long early intervals were largely 'wasted' instead of providing the intuitively expected advantage. That longer early anticipation intervals led to more time used seems best in-

terpreted as due to a performance factor, while the failure to utilize longer feedback frames early in training seems better attributed to a learning-to-learn problem. Prior practice on an unrelated verbal-discrimination task should overcome the latter and might also reduce the former problem. To some extent, the absence of a learning-to-learn confounding is implicit in the Cooper and Pantle (1967) 'rehearsal only' condition, but perhaps this source of confounding should be made explicit, since it could be involved in the past failures to confirm the total-time hypothesis in verbal-discrimination learning.

Those discrepancies from the hypothesis could also indicate that something other than the mere rehearsal of the R-term is occurring—the utilization of an association between the R- and W-term, for example, though that is quite speculative. The present matching data suggest that any such association was hindered when the early intervals were short, but since there was no significant difference in final performance levels, any causal influence from the association seems unlikely, although the same terminal level could have been reached in quite different ways. Obviously, the problem of temporal parameters in verbal-discrimination learning is a complex one, and it is only aggravated by any displacement of rehearsal to the anticipation frame.

Notes

This experiment was supported by Grant USPH 1 R03 18690-01 from the National Institute of Mental Health. Received for publication March 8, 1971.

References

- Cooper, E. H., and Pantle, A. J. 1967. The total-time hypothesis in verbal learning. *Psychological Bulletin* 68:221-234.
- Ekstrand, B. R., Wallace, W. P., and Underwood, B. J. 1966. A frequency theory of verbal-discrimination learning. *Psychological Review* 73:566-578.
- Farzanegan, F. 1969. Temporal variables in verbal discrimination learning. Unpublished doctoral dissertation, Saint Louis University. *University Microfilms* 69:16,022.
- Fulkerson, F. E., and Johnson, J. E. 1971. Methodological variables in verbal discrimination learning. *Psychonomic Science* 22:68-69.
- Ingison, L. J., and Ekstrand, B. R. 1970. Effects of study time, method of presentation, word frequency, and word abstractness on verbal discrimination learning. *Journal of Experimental Psychology* 85:249-254.
- Kanak, N. J. 1968. The effects of rate of exposure upon simultaneous intentional and incidental verbal-discrimination learning. *Psychonomic Science* 12:141-142.

- Mueller, J. H., and Slaymaker, F. L. 1970. Total time and stimulus-response imagery in paired-associate learning. *Journal of Experimental Psychology* 75:288-292.
- Nodine, C. F. 1969. Temporal variables in paired-associate learning: The law of contiguity revisited. *Psychological Review* 76:351-362.
- Nodine, C. F., and Goss, A. E. 1969. Temporal parameters in paired-associate learning. *Psychonomic Monograph Supplements* 3(1, whole no. 33).
- Radtke, R. C., McHewitt, E., and Jacoby, L. 1970. Number of alternatives and rate of presentation in verbal discrimination learning. *Journal of Experimental Psychology* 83:179-181.
- Skeen, D. R. 1970. Study interval, manipulated frequency, and verbal discrimination learning. *Psychological Record* 20:111-117.
- Wike, S. S., and Wike, E. L. 1970. Presentation rates and verbal discrimination learning. *Psychonomic Science* 19:253-254.
- Winer, B. J. 1962. *Statistical principles in experimental design*. New York: McGraw-Hill.
- Zacks, R. T. 1969. Invariance of total learning time under different conditions of practice. *Journal of Experimental Psychology* 82:441-447.

continued from page 62

- Lerner, Daniel. *Psychological Warfare against Nazi Germany: The Skyewar Campaign, D-Day to VE-Day*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xxv, 377. \$12.50; paperback, \$3.95.
- Liebert, Robert. *Radical and Militant Youth*. New York: Praeger Publishers, 1971. Pp. xvii, 257. \$7.95.
- Lord, Jesse R. *Marijuana and Personality Change*. Lexington, Mass.: Heath Lexington Books, 1971. Pp. xiii, 124. \$12.50.
- Lowry, Richard. *The Evolution of Psychological Theory: 1650 to the Present*. Chicago: Aldine, 1971. Pp. xviii, 237. \$8.95.
- McGuire, William, exec. ed. *The Collected Works of C. G. Jung*. Vol. 6. Bollingen Series, 20. Princeton, N.J.: Princeton University Press, 1971. Pp. xv, 617. \$15.00.
- Maddi, Salvatore R. *Perspectives on Personality: A Comparative Approach*. Boston: Little, Brown and Co., 1971. Pp. xi, 541. \$6.95.
- Margolis, Herbert F., and Rubenstein, Paul M. *The Groupsex Tapes*. New York: David McKay, 1971. Pp. vi, 306. \$6.95.
- Mayer, Morris F., and Blum, Arthur, eds. *Healing through Living: A Symposium on Residential Treatment*. Springfield, Ill.: Charles C Thomas, 1971. Pp. x, 215. \$15.50.
- Mittler, Peter. *The Study of Twins*. Baltimore: Penguin Books, 1971. Pp. 192. Paperback, \$2.25.
- Morrison, A., and McIntyre, D. *Schools and Socialization*. Baltimore: Penguin Books, 1971. Pp. 240. Paperback, \$1.65.
- Morrow, William R., ed. *Behavior Therapy Bibliography, 1950-1969*. Columbia: University of Missouri Press, 1971. Pp. 165. \$10.00.
- Munn, Norman L. *The Evolution of the Human Mind*. Boston: Houghton Mifflin, 1971. Pp. ix, 278. Paperback, \$4.50.
- Myklebust, Helmer R. *The Pupil Rating Scale: Screening for Learning Disabilities*. New York: Grune & Stratton, 1971. Pp. 40. Paperback, \$6.00.
- Noland, Robert L., ed. *Counseling Parents of the Ill and the Handicapped*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xxi, 606. \$15.75.
- Ord, I. G. *Mental Tests for Pre-literates*. London: Ginn & Co., 1971. Pp. xiii, 270. £2 75s.
- Paoletti, Rodolfo, and Davison, Alan N., eds. *Chemistry and Brain Development*. New York: Plenum Publishing, 1971. Pp. x, 457. \$22.50.
- Parker, Beulah. *My Language Is Me: Psychotherapy with a Disturbed Adolescent*. New York: Ballantine Books, 1971. Pp. x, 341. Paperback, \$1.25.
- Pease, Daniel C., ed. *Cellular Aspects of Neural Growth and Differentiation*. UCLA Forum in Medical Sciences, 14. Berkeley and Los Angeles: University of California Press, 1971. Pp. 520. \$25.00.
- Reisman, John M. *Toward the Integration of Psychotherapy*. New York: Wiley, 1971. Pp. xii, 155. \$9.50.
- Retterstol, Nils. *Long-term Prognosis after Attempted Suicide: A Personal Follow-up Examination*. Springfield, Ill.: Charles C Thomas, 1971. Pp. 112. \$11.00.
- Retterstol, Nils. *Prognosis in Paranoid Psychoses*. Springfield, Ill.: Charles C Thomas, 1971. Pp. 173. \$11.75.

continued on page 116

PROBABILITY LEARNING: CONTINGENT-EVENT SCHEDULES WITH LAGS

Richard B. Millward and Arthur S. Reber
Brown University and Brooklyn College, CUNY

A contingent-event schedule with lag L is defined so that $P(E_{1,n+L}|E_{i,n}) = \pi_{i1} \neq \bar{\pi}$, the marginal E_1 event probability. The 16 subjects were run at $L = 1, 3, 5$, and 7 over four days, with $\bar{\pi}$ at either .50 or .75 throughout and with both perseverance and alternation sequences given each day. The effect of (i.e., memory for) the contingent event decreased with lag: at $L = 1$, $P(A_{1,n+L}|E_{i,n})$ was slightly over π_{i1} , and at $L = 7$ it was very near $\bar{\pi}$. The subjects also responded to the frequency and pattern of noncontingent events in the sequences.

The importance of memory for the immediately preceding sequence of events has been a recurrent theme in the two-choice prediction experiment. The degree to which subjects attend to preceding events, the manner in which they process the information contained in the event sequence, and the role that the past events play in determining prediction responses are all issues of considerable theoretical importance.

There have been three basic ways to investigate the dependency between preceding events and prediction responses. The simplest and most indirect is statistical, as in the analyses of event patterns and recency curves (changes in response proportions as a function of the length of event runs). A second approach is to measure directly the subjects' memory for past events. A third approach is to incorporate dependencies directly into the event sequence itself and to measure the degree to which the subjects' response probabilities reflect those dependencies. Each method has implications for the theory of probability learning, as a brief review will show.

For example, the event-response contingencies revealed through the more general statistical analyses had considerable impact on the theories based on linear operators (Estes, 1959) and on Markov chain models (Myers and Atkinson, 1964). These models determine the event-response dependency by the effect the event sequence has on the state of conditioning. For linear operators, the state is the probability of an A_1 response on a given trial; for Markov chains, it is one of the states defined

by the Markov chain itself. The important point is that, given the state, the particular sequence of events that led to it does not influence the theoretical response probability. As Bush and Mostellar (1955) noted, such theories meet the independence-of-path assumption. But these models are basically memoryless, and the lack of a monotonic recency function is ample evidence against them (Anderson, 1960, 1964; Nicks, 1959).

An offshoot of the statistical method was the attempt to redefine the stimulus as a function of the past sequence of events. The definition most often assumed was that the subjects encode runs of events (Gambino and Myers, 1967; Restle, 1961, 1966; Rose and Vitz, 1966; Vitz and Todd, 1967), although other investigators (Anderson, 1964; Friedman, Carterette, and Anderson, 1968) emphasized the encoding of alternation of events or, more generally, assumed an encoding scheme that allows for all event patterns back five trials as stimuli (Feldman and Hanna, 1966). The theoretical questions raised by these approaches are what aspect of the past sequence acts as a stimulus, and how the stimulus-as-coded and the response are related. There has not yet been a theoretical redefinition of the stimulus that also handles most of the results observed in simple noncontingent probability learning.

The second general approach has been characterized by the attempt to directly measure memory for past events. Several studies (Millward and Reber, 1968; Reber and Millward, 1965; Vitz and Hazan, 1969) have asked subjects to recall directly some aspect of the past sequence of events. The results indicated good recall back at least five trials and better than chance recall back even further (perhaps as many as 14 trials). The important role that runs play in determining memory for past events was also illustrated (Millward and Reber, 1968) by the finding that memory decreased more rapidly as a function of event changes than as a function of trials. The same study also found that when subjects guessed about past events, they had a response bias that matched the event probabilities. So although the direct-measurement method has not been used to test theoretical developments, its results indicate that any new theory of probability learning will have to incorporate memory for past events and to maintain something approximating the matching rule.

The final approach has been to evaluate memory effects in probability learning by modifying the statistical structure of the event sequence. By varying the contingencies of events over various lags, the influence of some past trial can be measured fairly directly. There have been a number of studies in which the event sequence was generated by a Markov chain. In these studies the $P(E_{j,n+1}|E_{i,n}) = \pi_{ij}$ is defined so that $\pi_{11} \neq$

π_{21} . Anderson (1960) and Engler (1958) found that subjects' responses were related to the contingencies—that is, that the conditional response probabilities, $P(A_{1,n+1}|E_{i,n})$, $i = 1, 2$, were different. Although the conditional response probabilities were close to π_{i1} , they were generally above π_{i1} if $\pi_{i1} > .5$ and below π_{i1} if $\pi_{i1} < .5$. This is an 'overshooting' on the conditional response probabilities similar to the overshooting typically found for the marginal probability, $P(A_{1,n})$, in probability-learning experiments run with simple Bernoulli event sequences. Recently, Strub and Erikson (1968) investigated a contingency defined with a lag of two trials: $P(E_{k,n+2}|E_{j,n+1}, E_{i,n}) = \pi_{ik}$. In this sequence the $E_{k,n+2}$ event is independent of the $E_{j,n+1}$ event but dependent on the $E_{i,n}$ event. They found that response probabilities were influenced by the second-order dependencies but that the extent of the influence was determined by the degree of the dependency (the value of π_{ik}) and by the particular independent event on trial $n + 1$. They felt that their results were best explained by the Gambino and Myers (1967) generalization-of-runs model.

The present experiment was an extension of the Strub and Erikson (1968) study. Instead of defining the event contingencies over only two trials, we used lags of 1, 3, 5, and 7 trials. Further, we looked at a case where the marginal E_1 event probability was not .50. All of our subjects were highly practiced and all were run under every lag condition. They were brought to a stable level of responding by using the observation technique introduced by Reber and Millward (1968). Our intention was to make the contingency as salient as possible without telling the subjects of its presence. Thus, on the first day the contingency was defined with a lag of one trial, on the second day the lag was shifted to three trials, on the third day to five trials and on the fourth day to seven trials.

METHOD

—Subjects—Sixteen junior and senior high school students were each paid \$12.00 for eight hours of service. They were selected arbitrarily from a pool of students who answered an advertisement for subjects.

—Stimulus sequences—The main independent variable in this experiment was the probabilistic contingency between reinforcing events at different lags. Two conditional probabilities defined these contingencies for each of the four lag (L) conditions of the experiment. The conditional probabilities were

$$P(E_{1,n+L}|E_{i,n+L-1}, E_{j,n+L-2}, \dots, E_{1,n}) = \pi_{i1} \text{ and} \\ P(E_{1,n+L}|E_{i,n+L-1}, E_{j,n+L-2}, \dots, E_{2,n}) = \pi_{21},$$

for $i, j = 1, 2$; $L = 1, 3, 5, 7$. In order to use these probabilities to define a

sequence, the first L trials must be initially defined. This was done by generating independent Bernoulli trials using the overall average probability of an E_1 event, $\bar{\pi}$; where $\bar{\pi}$ is defined as $\pi_{21}/(\pi_{12} + \pi_{21})$. Certain characteristics of the sequences generated by these conditional probabilities should be emphasized. First, if the lag is L , then all sequences of length L ($E_n, E_{n+1}, \dots, E_{n+L-1}$) are independent and have frequencies generated by the average event probabilities. For example, when $L = 3$, then $P(E_{1,n+2}, E_{2,n+1}, E_{1,n}) = \bar{\pi}(1-\bar{\pi})\bar{\pi}$. If $\bar{\pi} = .5$, then all sequences of length L occur with equal frequencies. On the other hand, the sequences of length greater than L are not independent. Thus again, if $L = 3$, then the sequences of length 4, 5, and 6 show the dependencies

$$\begin{aligned} P(E_{l,n+3}, E_{k,n+2}, E_{j,n+1} | E_{k,n+2}, E_{j,n+1}, E_{i,n}) &= \pi_i, \\ P(E_{m,n+4}, E_{l,n+3}, E_{k,n+2} | E_{k,n+2}, E_{j,n+1}, E_{i,n}) &= \pi_i \pi_{jm}, \text{ and} \\ P(E_{o,n+5}, E_{m,n+4}, E_{l,n+3} | E_{k,n+2}, E_{j,n+1}, E_{i,n}) &= \pi_i \pi_{jm} \pi_{ko}, \end{aligned}$$

where l, m , and $o = 1, 2$ and $\pi_i = \bar{\pi}$ for $i = 1$ and $1 - \bar{\pi}$ for $i = 2$. All longer sequences will always show dependencies.

—Design—There were two groups of subjects, each group defined by the average probability of an E_1 event: group 1 with $\bar{\pi} = .50$, and group 2 with $\bar{\pi} = .75$. On days 1, 2, 3, and 4 the lags used in defining the event sequence were 1, 3, 5, and 7 respectively. We intentionally confounded practice with lag because we were primarily interested in what the subjects could do with practice and not what their performance might be in a short testing period. There were also two sequences within each group: sequences A and B. These were defined by the two conditional probabilities, π_{11} and π_{21} . Sequence A for group 1 had $\pi_{11} = .8$ and $\pi_{21} = .2$; group 2 had $\pi_{11} = .8$ and $\pi_{21} = .6$. Sequence B for group 1 had $\pi_{11} = .2$ and $\pi_{21} = .8$; group 2 had $\pi_{11} = .7$ and $\pi_{21} = .9$.

On each of the four days, each two-hour session was divided into two one-hour periods. Each period consisted of 1,000 observation trials, followed by 500 prediction trials generated according to a sequence A or B schedule. For days 1, 2, 3, and 4 the sequence types were ordered A-B, B-A, A-B, and B-A respectively. Thus, each subject received 8,000 observation trials and 4,000 prediction trials over the four days. All his trials had the same average event probability ($\bar{\pi} = .50$ or $.75$) but varied within each day in the sequence type (A or B) and across days in the lag ($L = 1, 2, 3$, and 4) of the contingency.

—Procedure—Observation trials are trials on which only events occur. No signal is presented and the subject does not predict which event he expects to occur. He is just told that he will see a series of events, to observe it, and that later he will have to predict the next event in a sequence probabilistically identical to the one he saw. In this experiment, the rate of presentation of the events on observation trials was two per second. (See Reber and Millward, 1968, for details of this procedure and its effectiveness as a training procedure.)

The subjects were read standard probability-learning instructions plus the instructions for the observation trials. The 1,000 observation trials took about 8.3 min. Then there was a short pause to tell the subjects that they were to

begin predicting. The 500 additional trials followed. These took about 5 sec per trial, or 42 min in all. The subjects were given a break after the first hour, before the second hour of running with the other sequence type. Four subjects were run at one time, with the same observation and prediction sequences. For each group, two subgroups of four subjects were run.

RESULTS

Average response proportions

Since each block of 500 prediction trials was preceded by 1,000 observation trials and since the subjects in each group were run with a constant average-reinforcement probability, $\bar{\pi}$, the average response proportions should be at a very stable asymptote. Figure 1 presents the asymptotic response proportions. They appear quite stable and are similar to those found in a number of other experiments (Reber and Millward, 1968, 1971). Group 1's overall average $P(A_1)$ equaled .513, and group 2 showed the usual overshooting, with a $P(A_1)$ of .787. In order to test for the effects of the two sequences and the four lags, an analysis of variance with one between-subjects and two within-subject factors was run for both groups. The only significant effect for group 1 was the difference between sequences [$F(1, 6) = 13.15, p < .025$]. The overall response proportion for sequence A was .525, and for sequence B .500. Although this is a small difference, it was consistent across all four lags. The reasons for the difference lie in an anomalous difference in the over-

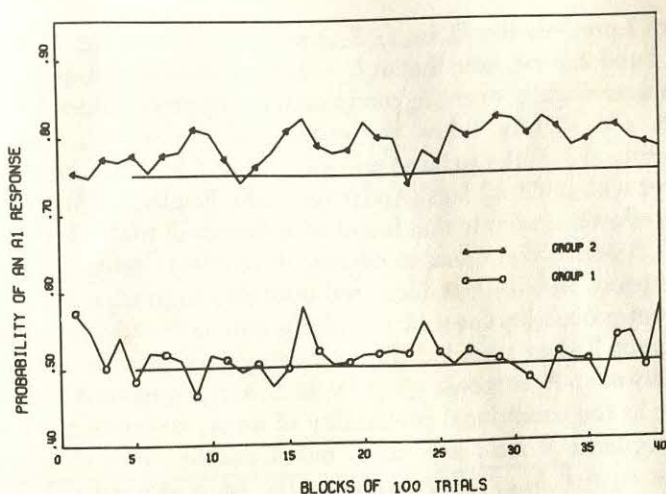


Fig. 1. Average response proportions for groups 1 and 2

all event frequencies for the two sequences. Sequence A had a proportion of .54 E_1 events, while sequence B had .50.

A similar analysis of variance for group 2 also yielded a single anomalous marginally significant interaction between the sequence and the lag variables [$F(3, 98) = 3.29, p < .025$]. This could be due to differential response probabilities for the two hours of each experimental session. The first hour of each session yielded an average response proportion across all four days of .78, and the second hour yielded an average of .80. Because the analysis of variance was run with A and B sequences as separate factors whereas their order had alternated, the interaction resulted. We have no explanation for a high $P(A_1)$ on the first hour of each session.

Figure 2 shows the individual subjects' response proportions plotted in blocks of 500 trials. (Note: the A_2 response proportions are plotted for group 2.) Some subjects seem to have responded systematically high or low, but others showed large shifts in their response levels. Since the response proportions were based on 500 trials, each of the eight points represents an individual subject's response proportion for a given condition of the experiment in this order: lag 1, sequences A, B; lag 3, sequences B, A; lag 5, sequences B, A; lag 7, sequences A, B. The analysis of variance failed to show a significant subjects effect and, therefore, the following analyses involve data pooled across subjects.

Memory effects

Figure 3 presents the $P(A_{1,n+L}|E_{j,n})$ as a function of lag, L , for both groups 1 and 2. First, note that at $L = 1$, the conditional response probabilities were slightly over the conditional event probabilities for $\pi_{11} = \pi_{21} = .8$, and slightly below the event probabilities $\pi_{11} = \pi_{21} = .2$. These results are similar to those previously found with contingent-event sequences with one-trial lags (Anderson, 1960; Engler, 1958). The overshooting effect is similar to that found with Bernoulli trials. However, for lags of 3, 5, and 7, there was no overshooting effect; instead, the conditional response probabilities decreased until they approximated the marginal event-probability level ($\bar{\pi} = .5$ for group 1, $\bar{\pi} = .75$ for group 2).

The group 1 data revealed the interesting fact that the conditional probability of an A_1 response given event E_i and sequence A was nearly the same as the conditional probability of an A_1 response given event E_j and sequence B for $i \neq j$; or, to put it another way, $P(A_{1,n+L}|E_{i,n} \text{ sequence A}) = P(A_{1,n+L}|E_{j,n} \text{ sequence B})$, for $i \neq j$; $i, j = 1, 2$. This means that, at least in the $\bar{\pi} = .5$ condition, an E_1 or an E_2 event could

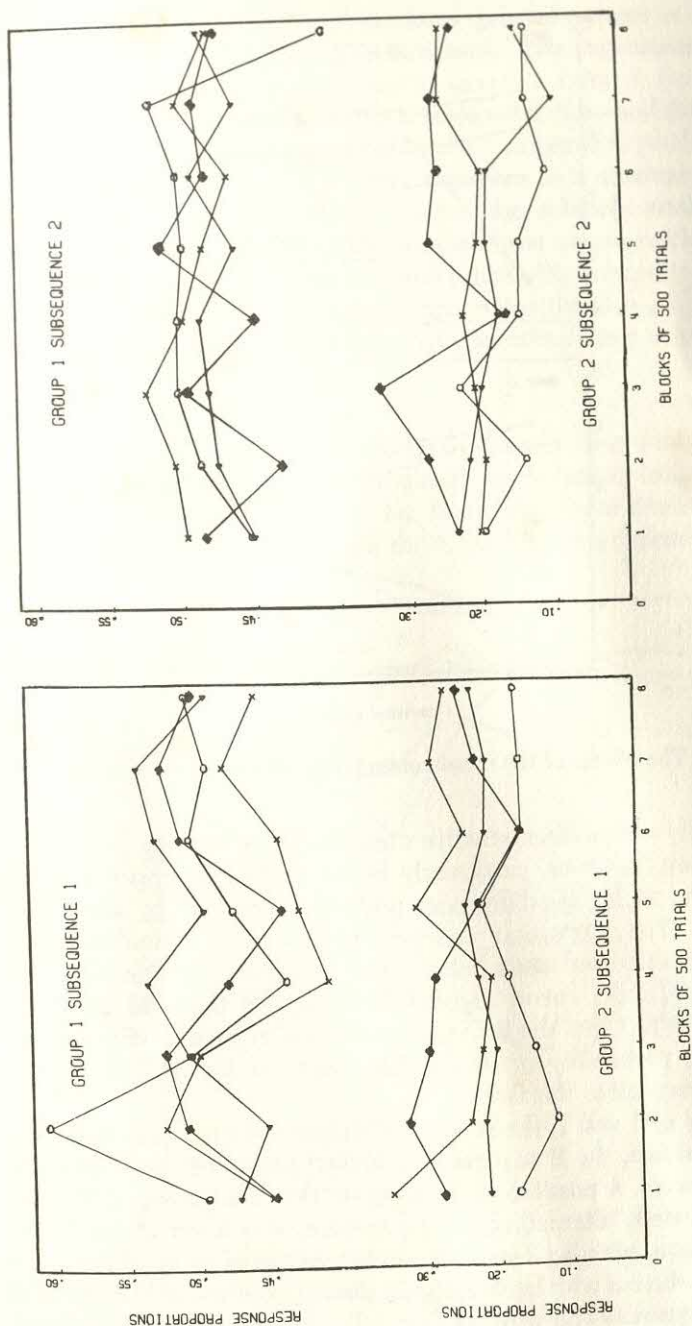


Fig. 2. Individual response proportions; each point represents a condition of the experiment; in the left and right panels, the squares are for subjects 4 and 8 respectively, the crosses for subjects 3 and 7, the triangles for subjects 2 and 6, the circles for subjects 1 and 5

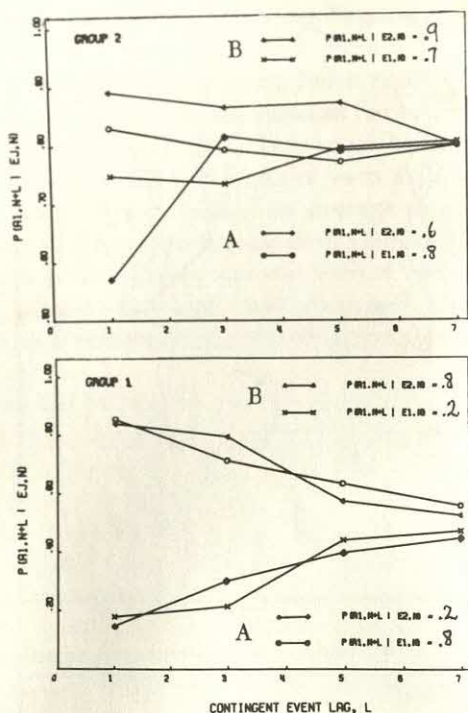


Fig. 3. The effects of the event contingency on response proportions

serve equally well as discriminative stimulus. There were some small differences here, however, particularly between the two upper and two lower curves at lag 3—differences probably produced by alternation tendencies. (The alternation tendency showed up in a number of statistics and is discussed again below.) The B sequences produced curves closer to 1.0 (upper curve) and 0.0 (lower curve) than did the A sequences. The fact that the B sequences tended to have shorter runs increased the probability of single alternations at lag 3. The subjects seemed to respond to this fact.

With lags of 5 and 7, the strong alternation tendency was no longer observed. In fact, the B sequences produced points nearer .5 than did the A sequences. A possible explanation for this change is that the proportion of strictly alternating event sequences was lower at lag 5 than at lag 3. That is, with lag 3 strict alternation occurred on $\frac{1}{4}$ of the event sequences, whereas with lag 5 strict alternation occurred only on $\frac{1}{16}$ of the event sequences and with lag 7 on only $\frac{1}{64}$ of the event sequences. (See the section on stimulus sequences, under Method, above.)

The group 2 results followed the same general pattern as those for group 1, but the effects are not as noticeable. The explanation for this may well be that when $\bar{\pi}$ decreases or increases from .5, the absolute probability of an E_1 event becomes a more valid cue—and that the degree to which a probabilistic event controls a response depends upon its validity. In other words, as $|\bar{\pi} - .5|$ increases in a contingent probability-learning experiment, it will control the subjects' prediction responses relatively more than other probabilistic contingencies. In this case, for lags greater than 1 the absolute probability seemed to dominate the subjects' predictions. The only exception was the $P(A_{1,n+L}|E_{2,n}) = .9$ curve, the only curve with a substantially greater degree of validity.

Recency curves

In this experiment, A sequences had relatively long runs and B sequences relatively short ones compared to Bernoulli trials with the same absolute $\bar{\pi}$ values. As the lag increased, these deviations from Bernoulli trials diminished, since the trials between dependent events

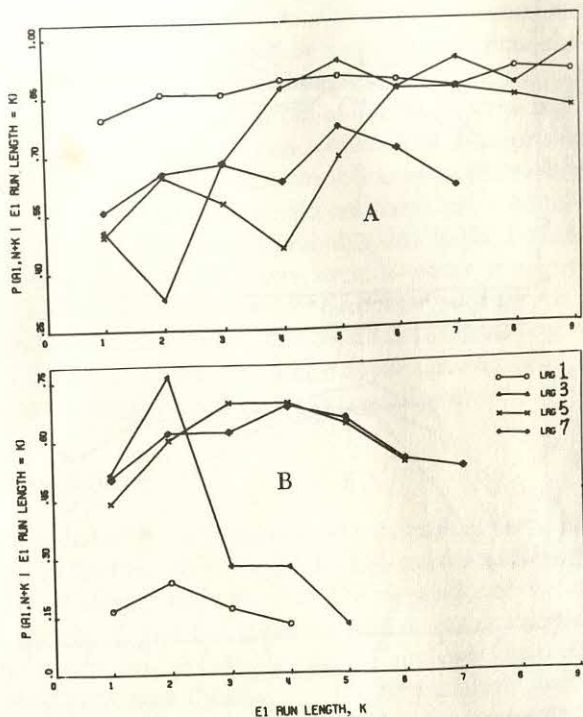


Fig. 4. Recency curves for the two sequence types for group 1

in a sequence were independent. Figures 4 and 5 present the recency curves for the two sequences of groups 1 and 2 respectively. The function plotted is $P(A_{1,n+k}|E_1 \text{ run of length } k)$. The $P(A_{1,n+k}|E_2 \text{ run of length } k)$ curves are not plotted for group 1 because they were very similar to those presented; that is, the $P(A_{2,n+k}|E_2 \text{ run})$ curve of sequence A (B) was similar to the $P(A_{1,n+k}|E_1 \text{ run})$ curve of sequence B (A). The group 2 frequencies for $P(A_{2,n+k}|E_2 \text{ run})$ were too low to be considered reliable, so these curves are also not plotted.

For group 1, sequence A, the contingent event is clearly evident in the runs analysis. For lag 1, the curve was high and fairly level for all run lengths. For lag 3, there was a nearly chance level of responding for $k = 1$. However, when the run length was two— $P(A_{1,n+3}|E_{1,n+2}, E_{1,n+1}, E_{2,n})$ —the curve dipped because the trial- n event (lag 3) was an E_2 and an E_2 thus tended to occur on trial $n + 3$ with probability .80. After a run of exactly three E_1 events, the probability of an A_1 response on trial $n + 4$ increased, since the E_1 event probability was .8 ($\pi_{11} = .8$). But note that this A_1 probability did not equal .8. There was still some effect of the E_2

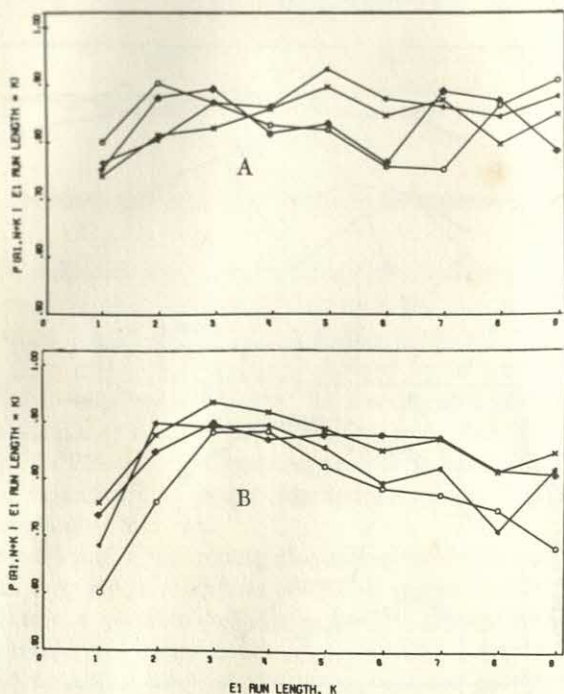


Fig. 5. Recency curves for the two sequence types for group 2

event back four trials. Then, after a run of length four and greater, the response probability was asymptotic at about .85. Thus, the fact that $P(A_{1,n+L}|E_{1,n})$ was over .85 was due in part to long runs of E_1 events. It was not a matter of isolating just the E_1 event back three trials and showing high response probability on that cue alone. The same pattern was found in the curve for lag 5, with essentially chance responding for $k = 1, 2$, and 3, a dip at $k = 4$, a rise at $k = 5$, and asymptotic responding for $k \geq 6$.

The B sequences for group 1 are interesting contrasts to the A sequences. The lag 1 and 3 curves for the B sequence are almost exact complements of those for the A sequence. (The curves are shorter because runs greater than the lag value were very rare in the B sequences.) But the lag 5 and 7 curves did not show any large effects of the contingencies with the B sequences. The discrepancy at lag 5 may be due to the different overall structures produced by the A and B sequences, or to the effects of previous experience with long runs. In particular, the subjects had just received two one-hour sessions with the A sequences with lags 3 and 5. Thus, because of this immediate history and because there were few runs of length five, the subjects may not have experienced enough runs of that length to learn that they were only continued one time in five. (In the earlier training, where long runs were frequent, runs of length five were continued in four out of five cases.) The lag 7 data were almost identical with the A-sequence data—not a surprising similarity in view of the Bernoulli characteristic of the seven intervening trials.

For group 2 the recency curves did not show any contingency effects even with lags 1 and 3. This was probably due to the dominance of the absolute probability, $\bar{\pi} = .75$. There were, however, some overall differences between the two sequences. In particular, sequence B produced a greater negative recency effect. This result is in agreement with that of Myers and Gambino (1966), who also found greater negative recency with sequences characterized by short runs of events.

Sequential analysis

It is known that subjects remember events back as far as five trials and maybe even as far back as seven trials (Millward and Reber, 1968). The recency curves clearly indicate that the dependency between the response on trial $n+L$ and the event on trial n was not entirely independent of the intervening events. Sequential analyses illustrate the nature of the dependency more thoroughly. The first analysis (individual-trials analysis) looks at the effect on $P(A_{1,n+k})$ response proportions of a given

event on trial n . The second analysis (pattern analysis) looks at the effects of all events together back k trials.

The pair of conditional probabilities looked at in the *individual-trials analysis*, the probabilities $P(A_{1,n+k}|E_{1,n})$ and $P(A_{1,n+k}|E_{2,n})$, show how an individual event on trial n influenced response probability on trial $n+k$, where k might or might not be the contingent event of the sequence. Tables 1 and 2 present these conditional probabilities for groups 1 and 2 respectively. Notice that the entries in the diagonal cells where $k = L$ are the probabilities plotted in Figures 3 and 4. The far-right column presents the overall proportion of A_1 responses for each group, sequence, and lag. The nonsignificant differences ($p \geq .05$) between E_1 and E_2 conditions (pairs of rows for a given lag) are indicated by an asterisk between the rows.

Table 1. Group 1's conditional response probabilities, $P(A_{1,n+k}|E_{1,n})$ and $P(A_{1,n+k}|E_{2,n})$

Lag	E_t	k							$P(A_{1,n})$	
		1	2	3	4	5	6	7		
Sequence A										
1	1	85	76	66	60					529
	2	15	26	37	44					
3	1	57	56	71	57	55				525
	2	47	48	30	48	50				
5	1	57	58	52	57	63	57			527
	2	48	47	54	47	40	48			
7	1	62	56	51	52	52	54	56		520
	2	41	47	53	52	52	49	47		
Sequence B										
1	1	18	71	36	60					506
	2	84	30	65	41					
3	1	52	50	21	47	49				498
	2	47	50	79	53	51				
5	1	54	61	49	50	44				508
	2	47	41	53	52	57				
7	1	57	55	49	50	47	48	47		497
	2	42	45	51	49	52	51	52		

Note: Decimal points omitted; the pair of probabilities above and below each asterisk are nonsignificant ($p > .05$).

Table 2. Group 2's conditional response probabilities, $P(A_{1,n+k}|E_{1,n})$ and $P(A_{1,n+k}|E_{2,n})$

Lag	E_i	k							$P(A_{1,n})$
		1	2	3	4	5	6	7	
Sequence A									
1	1	83	77 [*]	73	75				764
	2	57	76 [*]	85	81				
3	1	83	81	79	80	79			797
	2	69	76	82	80	82			
5	1	80	80	76	77 [*]	77	77 [*]		775
	2	72	73	81	79 [*]	79	78 [*]		
7	1	83	82	78	78	79 [*]	79	80	796
	2	71	73	84	83	80 [*]	82	80	
Sequence B									
1	1	75	84	78 [*]	79 [*]				785
	2	89	62	81 [*]	76 [*]				
3	1	81	80	74	75	75			772
	2	65	68	86	82	84			
5	1	85	84	80	80	79	79		812
	2	67	74	85	84	87	88		
7	1	83	83	79	79	80 [*]	80 [*]	80 [*]	802
	2	72	71	83	84	82 [*]	81 [*]	80 [*]	

Note: Decimal points omitted; the pair of probabilities above and below each asterisk are nonsignificant ($p > .05$).

The first fact to note is that there were relatively few nonsignificant differences. For $k = 1$ in group 1 all differences were significant, as they were for all $k = L$ conditions. A significant difference means that the probability of an A_1 response was differentially dependent upon the event that occurred k trials back. The size of the difference between the two values for a given lag indicates the magnitude of the differential effect. The pair showing the maximum difference in the direction of the contingency of each row is italicized. For the group 1 sequences, the maximum was clearly at $k = L$ except for $L = 7$. In general, the group 2 sequences had their maximum at $k = 1$ and did not show the strong effect of the contingent event.

The general pattern of results is that both the immediately preceding event and the contingent event have an effect on subjects' prediction re-

sponses. There was also a general trend for the effects of any one event to diminish as k increased, the trend being greater for B sequences than for A sequences. The fact that the contingencies defining the B sequences were opposites to those defining the A sequences (i.e., that E_1 events follow E_2 events and E_2 events tend to follow E_1 events with the B sequences) shows up in the conditional probabilities at the $k = L$ values. Note that $P(A_{1,n+L}|E_{1,n}) > P(A_{1,n+L}|E_{2,n})$ except at $k = L$ and in a few cases where k was near L . Of course, this pattern simply reflects the fact that the subjects were detecting the contingencies in the sequences, since the B sequences require this reversal. However, the reversal did not carry over to all k values, showing that, except at $k = L$, the E_1 events had their usual effect, which is to increase the probability of an A_1 response.

Now for the *pattern analysis*. The recency analysis indicated that the subjects were not simply isolating the contingent event back L trials and responding to it as an isolated cue. The *joint* probabilities $P(A_{1,n+L}, E_{i,n+L-1}, E_{j,n+L-2}, \dots, E_{k,n})$ provide further information about the effects of trials before and after the contingent event on trial n . However, the number of joint probabilities was extremely large and for illustrative purposes only the *conditional* probabilities for lag 3 are presented. (For the other joint probabilities, see Millward and Reber, 1971.)

Table 3 presents these conditional probabilities $P(A_{1,n+L}|E_{i,n+L-1}, \dots, E_{k,n}, E_{l,n-1}, E_{m,n-2})$ for lag 3 and extends the event sequence back two trials beyond the contingent event on trial n . Thus, it is also possible to see what effect events more remote than the contingent event had on the response probability. Table 3 also presents the conditional probabilities so that successive events starting with $E_{m,n-2}$ are collapsed ($E_{m,n-2}, E_{k,n-1}, E_{k,n}, \dots$).

The data in Table 3 permit a number of generalizations about this experiment. First, the contingent event clearly had the primary effect on subjects' predictions. Second, other statistically independent events also played a role in determining subjects' predictions. The effects of these independent events seem to have been determined largely by two factors: (a) the frequency of each event in the sequence and (b) the saliency of particular patterns of events.

A comparison of Table 3's first two columns indicates the effect of the critical event on trial $n-2$ with E_1 on trial $n-1$ and the fourth and fifth columns indicate the effect of the critical event on trial $n-2$ with E_2 on trial $n-1$. The similarities between the first and second and between the fourth and fifth columns indicate the effect of events more remote than the critical event; if these pairs of columns were identical it would mean

Table 3. Conditional response probabilities for group 1, lag 3, $P(A_{1,n+3}|E_{i,n+2}, E_{j,n+1}, E_{k,n}, E_{l,n-1}, E_{m,n-2})$

			$E_{m,n-2} = 1$	2	•	1	2	•	•
			$E_{l,n-1} = 1$	1	1	2	2	2	•
$E_{i,n+2}$	$E_{j,n+1}$	$E_{k,n}$							
Sequence A									
1	1	1	91	87	90	65	79	68	86
1	2	1	71	70	71	42	31	33	51
2	1	1	46	42	45	75	51	71	65
2	2	1	40	68	62	70	69	69	68
1	1	2	29	47	33	33	46	37	33
1	2	2	46	30	34	42	52	49	37
2	1	2	53	38	49	27	38	28	32
2	2	2	50	32	33	09	06	07	13
Collapsed			62	51	50	45	57	47	52
Sequence B									
1	1	1	13	27	25	25	29	28	28
1	2	1	33	30	33	14	18	14	19
2	1	1	11	12	12	39	16	23	14
2	2	1	17	34	21	46	58	48	25
1	1	2	80	65	67	44	88	81	77
1	2	2	77	50	74	88	84	88	85
2	1	2	81	92	89	67	68	68	85
2	2	2	72	65	71	74	65	72	71
Collapsed			53	52	46	48	52	47	50

Note: Decimal points omitted; conditional probabilities computed after collapsing over events are indicated above by centered dots; the seven patterns across the collapsed rows were (11•••), (12•••), (21•••), (22•••), (1••••), (2••••), and (•••••), the 5-tuple standing for $(E_{n+2}, E_{n+1}, E_n, E_{n-1}, E_{n-2})$.

that the more remote events $E_{i,n-1}$ and $E_{i,n-2}$ contributed nothing to the subjects' prediction behavior. The differences can be traced to event frequency and/or event patterns. In sequence A, for example, the 112 row has probabilities .29 and .47 in the first and second columns and probabilities .33 and .46 in the fourth and fifth columns. The .29 and .33 values seem fairly low—apparently as the result of an alternation tendency, since the two complete sequences were 11211 versus 21211 and 12211 versus 22211. In the former pair, it may be inferred, the subjects were predicting a 2 in order to complete the sequence 11211; they showed no such strong tendency for the 21211 sequence. Similarly, in the latter pair, 12211 produced a strong tendency to predict a 2 and

22211 did not. If one looks at all the fairly different pairs, 'simple' patterns seem to account for the differences.

However, it does not seem possible to account for all differences simply by referring to salient patterns; frequency was also an important factor. For example, when we looked at the pair of events that preceded the critical event, we found that the four conditional probabilities $P(A_{1,n+L} | E_{l,n-1}, E_{m,n-2})$ for $l, m = 1, 2$ were .59, .54, .50, and .44. These four probabilities seem to indicate an effect due simply to frequency, since the more E_1 events, the higher the A_1 probability.

Similar patterns can be found for the events between the critical event and the prediction response. The third and sixth columns in Table 3 (events on trials $n-2$ and $n-1$ collapsed) show the effect of the pattern of events between the contingent event and the response. The highest probability was when all events were 1's (the first row). The lowest probabilities occurred when all events were 2's (the last row). This could be due to frequency or to a tendency to respond to runs. The low values for pattern 121 in row 2 illustrate an alternation tendency, since row 3 has a higher probability with the E_2 event on the previous trial. And when we look at the collapsed conditional probabilities $P(A_{1,n+3} | E_{i,n+2}, E_{j,n-1})$, $i, j = 1, 2$, we see the probabilities .62, .51, .50, and .45, which are extremely close to the .59, .54, .50, and .44 (presented above) for the same four patterns two trials farther back. In general, this pattern of results is seen in complementary form in the data from sequence B.

DISCUSSION

Two broad generalizations can be made here. First, the memory for contingent events is evident back at least five trials and the effect of the contingent event is a decreasing function of the length of the lag. Second, the contingent-event effect is by no means uninfluenced by other events before and after the contingent event. There appears to be both a frequency effect and a pattern effect. The frequency effect is indicated because the response probability is related to the actual number of E_1 and E_2 events that precede the response. The pattern effect is indicated because for specific patterns, such as single and double alternations, the response probability varies as if the subject expected the pattern to continue.

The results do not tell us how past events are encoded, other than that the encoding process is differentially sensitive to contingencies built into the sequence. They are in very good agreement with the earlier work (Reber and Millward, 1965; Millward and Reber, 1968) on direct

measurement of memory for past events, however. The conclusion seems to be that memory is a decreasing function of trials back about five to seven trials and that it is affected to some degree by run structure, as observed directly by Millward and Reber (1968) and as observed here by the divergence of the curves in Figure 3 at lag 3. The results agree too with those of Strub and Erickson (1968), who also found that the contingency was strongly influential but that subjects did not respond independently of the events on intermediate trials.

The results about memory for past events in probability learning now seem complete and may be helpful in determining a better theoretical structure. It is obvious that any theory will have to incorporate memory for the last five to seven trials and that it will have to allow for this memory to have adjustable weights as a function of the contingencies involved in the sequence. The theory will have to state how these weights are determined as a function of the sequence. The weights will not be constant for all sequence types and will depend upon the absolute probability as well as contingency probabilities. Our hunch on this matter is that the subject can adjust the weights so that he attends to the most reliable or valid cues in the situation. When the absolute probability, $\bar{\pi}$, is high, other sequential dependencies become less useful to him than the absolute probability itself, upon which he will then heavily rely. To date we have not been able to specify any mechanism by which he accomplishes this adjustment of weights, although it seems clear that such a mechanism must be based on a coding system involving estimates of cue validity.

Perhaps we should state here that it does not appear that we are studying a simple conditioning phenomenon. Rather, experiments of the form traditionally classified as probability learning are experiments on complex information integration. Any future theory of probability learning must be a theory that involves memory functions, encoding mechanisms, and hypothesis behavior.

Notes

This research was supported by Grant MH11255 from the National Institutes of Mental Health. Received for publication February 16, 1971.

References

- Anderson, N. H. 1960. Effect of first-order conditional probability in a two-choice learning situation. *Journal of Experimental Psychology* 59:73-93.
- Anderson, N. H. 1964. An evaluation of stimulus sampling theory. In *Cate-*

- gories of human learning, ed. A. W. Melton, pp. 129-144. New York: Academic Press.
- Bush, R. R., and Mosteller, F. 1955. *Stochastic models for learning*. New York: Wiley.
- Engler, J. 1958. Marginal and conditional stimulus and response probabilities in verbal conditioning. *Journal of Experimental Psychology* 55:303-317.
- Estes, W. K. 1959. The statistical approach to learning theory. In *Psychology: A study of a science*, Vol. 2, ed. S. Koch. New York: McGraw-Hill.
- Feldman, J., and Hanna, J. F. 1966. The structure of responses to a sequence of binary events. *Journal of Mathematical Psychology* 3:371-387.
- Friedman, M. P., Carterette, E. C., and Anderson, N. H. 1968. Long-term probability learning with a random schedule of reinforcement. *Journal of Experimental Psychology* 78:442-455.
- Gambino, B., and Myers, J. L. 1967. Role of event runs in probability learning. *Psychological Review* 74:410-419.
- Millward, R. B., and Reber, A. S. 1968. Event-recall in probability learning. *Journal of Verbal Learning and Verbal Behavior* 7:980-989.
- Millward, R. B., and Reber, A. S. 1971. Probability learning: Contingent-event schedules with lag. Brown University Psychology Department Studies in Human Learning, 7.
- Myers, J. L., and Atkinson, R. C. 1964. Choice behavior and reward structure. *Journal of Mathematical Psychology* 1:170-203.
- Nicks, D. C. 1959. Prediction of sequential two-choice decisions from event runs. *Journal of Experimental Psychology* 57:105-114.
- Reber, A. S., and Millward, R. B. 1965. Event memory in probability learning. *Psychonomic Science* 3:431-432.
- Reber, A. S., and Millward, R. B. 1968. Event observation in probability learning. *Journal of Experimental Psychology* 77:317-327.
- Reber, A. S., and Millward, R. B. 1971. Event tracking in probability learning. *American Journal of Psychology* 84:85-99.
- Restle, F. 1961. *Psychology of judgment and choice: A theoretical essay*. New York: Wiley.
- Restle, F. 1966. Run structure and probability learning: Disproof of Restle's model. *Journal of Experimental Psychology* 72:382-389.
- Rose, R. M., and Vitz, P. C. 1966. The role of runs in probability learning. *Journal of Experimental Psychology* 72:751-760.
- Strub, M. H., and Erickson, J. R. 1968. Second-order dependency in probability learning. *Journal of Experimental Psychology* 78:261-268.
- Vitz, P. C., and Hazan, D. N. 1969. Memory during probability learning. *Journal of Experimental Psychology* 80:52-58.
- Vitz, P. C., and Todd, T. C. 1967. A model for simple repeated binary patterns. *Journal of Experimental Psychology* 75:108-117.

PATTERNS IN RANDOM CHOICE

William F. Orr, Charles F. Federspiel, and O. Dale Maxwell
Vanderbilt University

The 82 subjects who participated twice in sham ESP experiments showed markedly similar responses to card position: a deficiency in calls of symbols in extreme lateral positions, a deficiency in repetition of identical calls, and a tendency to include all five symbols in any five consecutive calls. Those responses did not vary significantly from one test period to another, nor with alteration of symbol, yet there were also unique subject variations.

Men have long understood that the likelihood of an event or sequence of events occurring varies between what seems certain and what seems chaotic and unpredictable. For more than 300 years a field of formalized probability theory, begun with the Pascal-Fermat correspondence, has been the subject of lively study, and domains of mathematical probability are now well delineated (see Nagel, 1955, p. 343; Reichenbach, 1949, p. 3).

Men also recognize that success in everyday life depends on the accuracy of their estimates of the probable occurrence of events. Jean Piaget (1951) was so impressed with the importance of the development of a concept of chance and probability in the child that he made this the topic of one of his major works.

Formal study of the *human concept* of chance is a recent comer. In 1938, Goodfellow analyzed the data from the Zenith radio experiments on 'telepathy.' He showed that the radio audience collectively used consistent and predictable methods to order their sequence of calls: when one of the 'preferred' sequences was presented to the audience, high levels of concordance between actual and audience-guessed sequences were observed; contrariwise, when an unlikely chosen sequence was presented, low concordance. In disproving the effect of telepathy, he examined the frequency with which (binary) sequences were called, factors influencing the first call of a sequence, and the influence of the preceding on the succeeding call (Goodfellow, 1938, 1940).

In asking for 'random' responses from individuals, statisticians find predictable patterns of behavior: "Experience has shown that the human being is an extremely poor instrument for conduct of a random selection.

Wherever there is any scope for personal choice or judgment on the part of the observer, bias is almost certain to creep in. Nor is this a quality that can be removed by conscious effort or training. Nearly every human being has as a part of his psychological makeup a tendency away from true randomness in his choices" (Yule and Kendall, 1950, p. 373). Later, Kendall and Stuart (1963, pp. 208-211) specified two of the nonrandom aspects: a general tendency to avoid calling extreme values from an admissible set, and a deficiency of serial repetition of identical calls.

Jenkins and Cunningham (1949) found that subjects attempting to randomize a series of ten digits called predictable sequences. They found that a given number was followed 25% of the time by the number one less than the index number and that the index number rarely succeeded itself. Others (especially Senders, 1953; Ross, 1955; Bakan, 1960; and Rath, 1966) have also reported studies of certain aspects of the concept of probability. Here we present a more comprehensive study of the human's idea of randomness.

We used the format of the card-matching ESP experiment of Rhine, Pratt, Stuart, Smith, and Greenwood (1940, p. 25, also Figure 2) to study the subjects' predictions. Investigations into their sequences of calls served as a method of exploring the processes involved in human choice. Through this we (a) tested the generality of the statements on human 'random choice' behavior posited above, (b) studied the stability of this behavior in time, (c) investigated some variables that may influence these choices, and (d) examined the subjects' strategies for patterning and ordering their responses.

METHOD

Eighty-two adult subjects were studied, 48% of them college graduates (medical students, professional, technical, and administrative personnel in the medical center), 52% of them in occupational classifications generally defined as nonprofessional. It must be emphasized that these 82 individuals represent only a sample of people in the medical center who were willing to devote the time required to perform the task demanded by the experiment. Of the 89 subjects who originally consented to participate, 7 could not be scheduled for a repeat trial and were excluded.

The experiment followed the conventional card-guessing ESP test and was carried out in a quiet room without windows. The subject was seated on one side of a black screen with a 24-by-3-in. section open at the bottom. On the other side of the screen, the experimenter was seated so that only his hands were visible. Five cards, on which were printed various symbols, were placed in a set order in the opening at the bottom of the screen in full view of both experimenter and subject. A test stack of 25 unscripted cards other-

wise identical in size and color was placed beside the experimenter. The subject was told:

This deck [indicating the test stack] consists of five each of the five symbols you see before you, for a total of 25 cards. Indicate by touch the symbol which you feel corresponds to the symbol on the top of this stack [again indicating the unscribed cards]. I will remove the top card after your choice. Then, touch the symbol which you feel corresponds to the symbol currently occupying the top position—the one that started in the second position—and we will proceed in this fashion through the 25 cards.

The examiner recorded (unseen by the subjects) the number position in front of the subject (rather than the symbol on the card) as his response. This response was referred to as a 'call.' The sequence of 25 consecutive calls obtained by displacing each of the 25 cards of the test deck constituted a 'unit.' For each set of card types, four units were carried out. Then the subject was presented with another group of five cards on which there were different symbols, and the test was repeated. Four units of 25 cards each were completed using each of the three card types. Cards of a given type constituted a 'deck,' and 100 calls from each of the three decks were elicited at a given sitting. The symbols of deck 1 were numerals 1 through 5, deck 2 were 'open' symbols, and on deck 3 'closed' symbols (see Figure 1). After an interval of no less than four weeks, each subject repeated the entire experiment with the order of the exposed cards unchanged and with identical instructions.

The subjects were not aware that there were no symbols on the cards in the experimenter's stack; that is, they were not told that this was a *sham* ESP experiment.

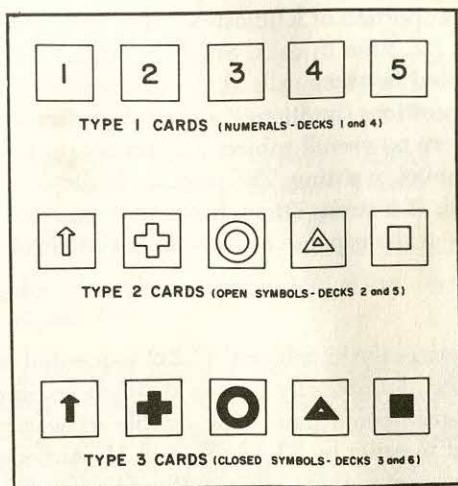


Fig. 1. Symbols and their positions

RESULTS

We have no basis for assuming that our subjects attempted to randomize responses. However, patterns and biases are demonstrated more clearly by comparing responses with those that would be expected if the responses were distributed equally among all possible alternatives. To do that, we first consider the effects of symbol variation and passage of time on single responses. Next, we take up doubles (two sequential responses), then triples, and finally some aspects of longer sequences.

Singles

The positional distribution of 49,200 single responses of the 82 subjects is shown by deck in Table 1. Although this table does not allow one to assess subject variation, it points up the gross and consistent patterns that prevailed regardless of sitting (time) or symbol type.

Each subject made a total of 600 calls, of which three-fifths, or 360, would be expected to be in positions 2, 3, or 4 if the calling pattern were uniform. For any given deck, each subject made 100 calls. Under uniformity, 20 calls would be expected for each of the five symbols. Figure 2 shows the observed and expected numbers of subjects making various numbers of midposition calls.

Detailed chi-square analyses (available from the authors) verified that there were no significant differences (either within-subject or among-subjects) in the proportion of leftmost-position calls (position 1) relative to right-position (position 5) calls. Similarly, there were no significant differences revealed between calls at the midposition (position 3) relative to adjacent positions (positions 2 and 4). Significance tests also verified that there were no overall subject differences (in positions selected) dependent on symbol or sitting. The primary finding from the examination of single calls is a strong (though not universal) tendency to make midposition calls at the expense of extreme-position calls.

Doubles

Table 2 summarizes the 82 subjects' 47,232 sequential responses of two calls (doubles) and dramatically reveals deficiencies in calling each of the possible like-membered pairs. The double 44 was called least frequently, followed in order by 33, 22, 55, and 11. Although the aversion to extremes is not seen in these like-membered pairs, the next least frequently called doubles were 15 and 51, an observation consistent with the findings reported for singles.

Table 1. Singles called for the various decks; times called

		Position				
		1	2	3	4	5
First sitting						
Deck 1 (numerals)	Expected	1,640	1,640	1,640	1,640	1,640
	Observed	1,495	1,761	1,705	1,708	1,531
	0-E	-145	121	65	68	-109
	Rank order	5	1	3	2	4
Deck 2 (open symbols)	Expected	1,640	1,640	1,640	1,640	1,640
	Observed	1,524	1,756	1,673	1,719	1,526
	0-E	-116	116	33	79	-114
	Rank order	5	1	3	2	4
Deck 3 (closed symbols)	Expected	1,640	1,640	1,640	1,640	1,640
	Observed	1,515	1,762	1,688	1,704	1,531
	0-E	-125	122	48	64	-109
	Rank order	5	1	3	2	4
Second sitting						
Deck 4 (numerals)	Expected	1,640	1,640	1,640	1,640	1,640
	Observed	1,501	1,702	1,734	1,733	1,530
	0-E	-139	62	94	93	-110
	Rank order	5	3	1	2	4
Deck 5 (open symbols)	Expected	1,640	1,640	1,640	1,640	1,640
	Observed	1,517	1,749	1,712	1,712	1,510
	0-E	-123	109	72	72	-130
	Rank order	4	1	2.5	2.5	5
Deck 6 (closed symbols)	Expected	1,640	1,640	1,640	1,640	1,640
	Observed	1,468	1,761	1,716	1,716	1,538
	0-E	-172	121	76	76	-102
	Rank order	5	1	2.5	2.5	4
Total (all decks)	Expected	9,840	9,840	9,840	9,840	9,840
	Observed	9,020	10,491	10,228	10,292	9,166
	0-E	-820	651	388	452	-674
	Rank order	5	1	3	2	4

Note: Total number of singles called, 49,200; in each deck, 8,200 (82 subjects $\times$ 100 calls in each deck).

A chi-square analysis of doubles (available from the authors) taking account of subjects, symbols, and sittings was carried out to assess the significance of the data presented in Table 2, as well as the effect of subject and deck differences. The difference between like- and unlike-membered doubles was, of course, strikingly significant. The relative frequency of calls for doubles representing midpositions (positions 2, 3, and 4) was significantly greater than for calls representing extreme posi-

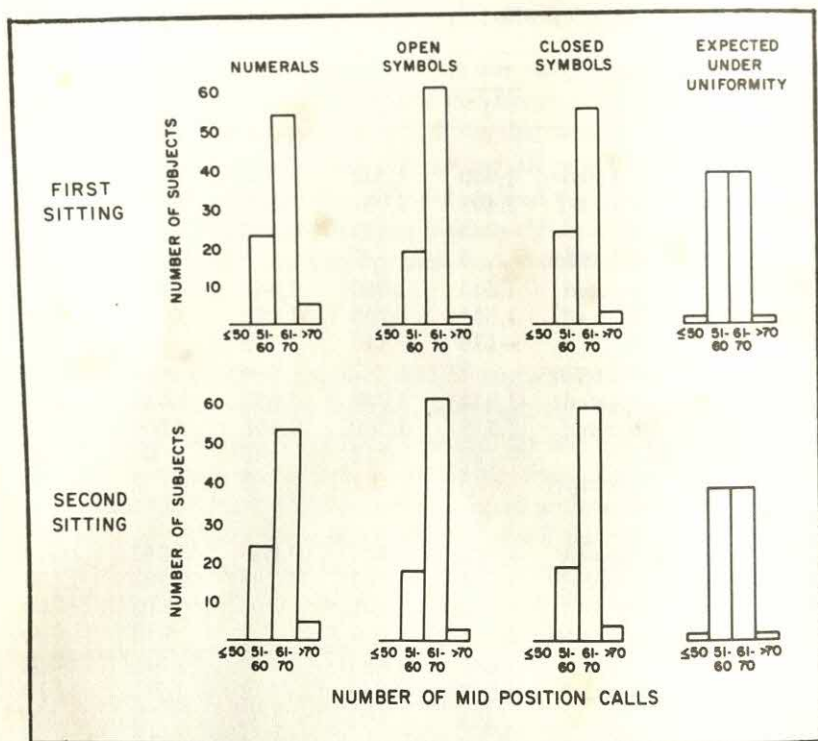


Fig. 2. Numbers of subjects making various numbers of midposition calls

tions; for example, the double 23 was called with relative frequency 5.4%. Similarly, double 43 was called with frequency 5.4%, and double 54 with frequency 4.8%. Statistical treatment of the data also showed significant differences between like- and unlike-membered pairs for both extreme and midposition calls.

In considering unlike-membered pairs, we may also observe that a slight preference for ascending doubles relative to descending doubles was demonstrated. By referring to Table 2 one can calculate the cumulative percentage for ascending doubles (12, 13, 14, 15, 23, 24, 25, 34, 35, 45) to be 47.8%; that for descending doubles (21, 31, 32, 41, 42, 43, 51, 52, 53, 54), 47.1%.

Although all 82 subjects exhibited a preference for unlike-membered doubles, there was considerable subject variation in the degree of preference for doubles composed of extreme values (positions 1 and 5) relative

Table 2. Doubles (sequential pairs) called for all decks

Positions of double	Times called	Frequency (%)	Positions of double	Times called	Frequency (%)
11	499	1.0	41	1,853	3.9
12	2,354	5.0	42	2,623	5.6
13	2,186	4.6	43	2,561	5.4
14	2,007	4.2	44	388	.8
15	1,578	3.3	45	2,526	5.3
21	2,431	5.1	51	1,747	3.7
22	480	1.0	52	2,218	4.7
23	2,557	5.4	53	2,076	4.4
24	2,688	5.7	54	2,252	4.8
25	1,912	4.0	55	484	1.0
31	2,139	4.5			
32	2,363	5.0			
33	429	.9			
34	2,567	5.4			
35	2,314	4.9			

Note: Total number of doubles called, 47,232 (82 subjects $\times$ 96 calls in each deck $\times$ 6 decks).

to doubles composed exclusively of midvalues (positions 2, 3, and 4). The group as a whole tended to choose ascending doubles rather than descending doubles. Here only small subject differences were seen. Of the subjects, 47 showed a slight preference for ascending doubles, and the remaining 35 showed very slight preferences for descending doubles. The different decks themselves produced no statistically significant differences in doubles called by the 82 subjects.

Triples

Examination of triples is restricted here to considering hypotheses relevant to order only. Table 3 lists the total number of calls of each of the different triples. Under uniformity, each of the 125 possible triples would be called 362 times. Triple 111 was in fact called 31 times and triple 423 was called 898 times.

Analysis is simplified by classifying the 125 different triples into nine groups according to their order characteristics. The nine groups were: static (111, 222, 333, 444, 555); ascending (the 10 triples exemplified by positional calls 124 and 235); descending (the 10 triples exemplified by 531 and 421); convex (the 30 triples exemplified by 131

Table 3. Triples called for all decks; times called

Positions of second and third members of triple	Position of first member of triple				
	1	2	3	4	5
11	31	163	104	101	88
12	163	299	671	555	562
13	120	785	143	546	451
14	86	672	616	116	463
15	83	411	494	460	95
21	270	115	648	673	625
22	135	34	106	116	103
23	616	118	184	898	589
24	728	120	800	142	742
25	503	92	527	665	77
31	121	805	78	575	489
32	655	187	116	721	565
33	72	120	38	98	95
34	627	758	109	202	724
35	605	569	73	847	129
41	82	648	541	49	488
42	679	186	804	116	689
43	578	880	220	87	670
44	63	134	118	31	75
45	533	755	787	94	259
51	70	475	602	480	70
52	460	119	686	797	87
53	417	562	188	692	106
54	469	576	665	277	146
55	81	85	94	150	62

Note: Total number of triples called, 45,264 (82 subjects $\times$ 92 calls in each deck $\times$ 6 decks).

and 354); concave (the 30 triples exemplified by 514 and 425); static ascending (the 10 triples exemplified by 335); static descending (the 10 triples exemplified by 221); ascending static (the 10 triples exemplified by 233); and descending static (the 10 triples exemplified by 411).

Considering the number of different triples in each class and the total number called, one can calculate the expected number for any class under uniformity. Table 4 shows the number and range of observed calls of various classes and the number expected under uniformity. Statistical analyses revealed highly significant differences between static

Table 4. Triples called for all decks; analyzed by class

	Static	Ascending	Descending	Concave	Convex
Expected	22.08	44.16	44.16	132.48	132.48
Observed	2.39	79.04	74.91	172.24	173.06
Range of individual calls	0-44	38-159	10-119	75-229	81-225
	Static ascending	Static descending	Ascending static	Descending static	
Expected	44.16	44.16	44.16	44.16	
Observed	12.90	11.83	12.83	12.79	
Range of individual calls	0-50	0-46	0-58	0-55	

and nonstatic triples, just as with doubles. But unlike doubles, the differences between ascending and descending triples were highly significant.

The analysis, for each of the 82 subjects, was on 92 triples for each of the six decks tested, or 552 triples per subject. All the subjects except one called fewer than the expected number of static (like-membered) triples. In fact, 55 (67%) of the 82 subjects chose no like-membered triples at all. The single subject who called more than the expected number of like-membered triples made 44 such calls. These calls consisted of two 111s, seven 222s, four 333s, eight 444s, and twenty-three 555s.

The highly significant group preference for ascending versus descending triples was not universal among the subjects but appeared to be due to the much larger number of ascending than descending calls in a relatively few subjects. There were 43 subjects who called more ascending than descending triples and 34 subjects who called more descending than ascending triples. (In 5 of the 82 subjects there was no difference between the number of ascending and descending triples called.) The 43 subjects who preferred ascending triples called 3,890 such triples, or 90.5 ascending triples per subject. The 34 subjects who preferred descending triples called an average of only 79.8 such triples per subject. No other differences in the classes of triples called were statistically significant, either for the composite group or among subjects.

Longer sequences

When we consider sequences of more than three calls, a priori analysis becomes cumbersome. Of the immense number of possible approaches, the following a posteriori hypotheses are suggested.

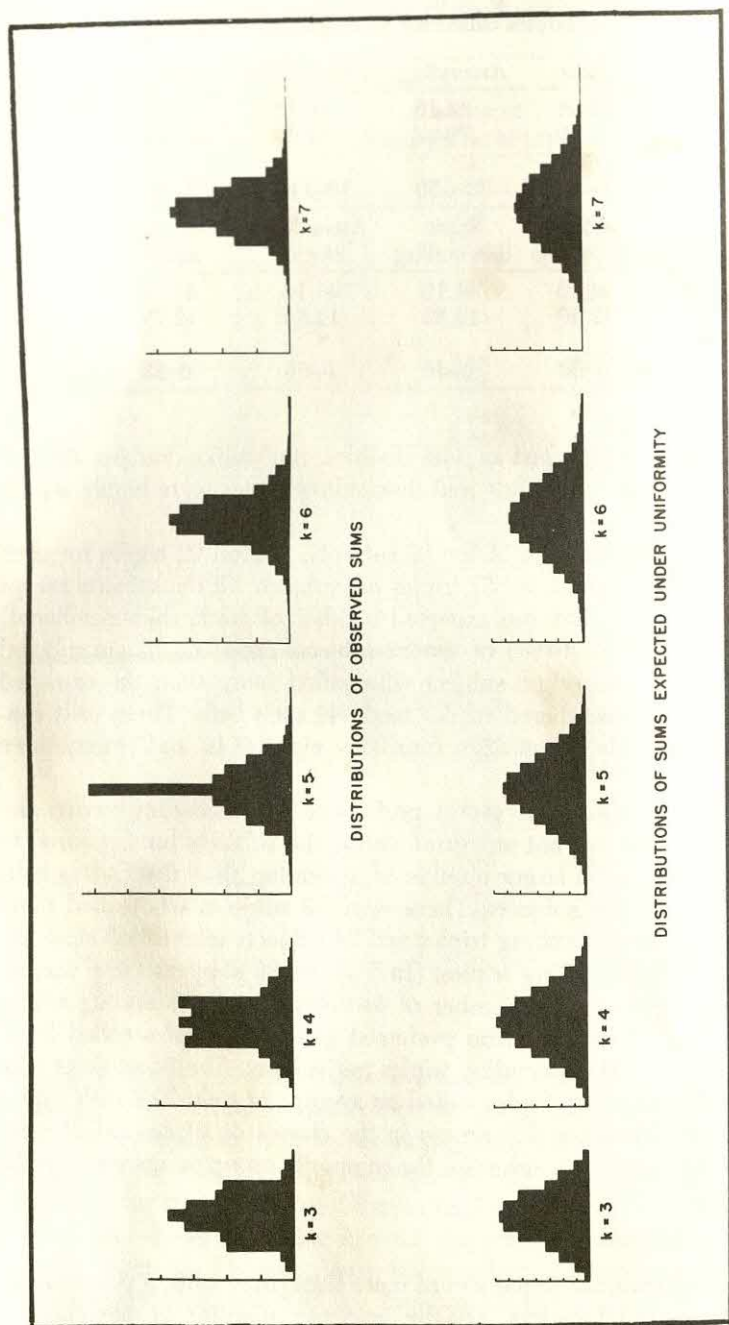


Fig. 3. Distributions of observed and expected sums of k consecutive calls

The raw data and previous work on sequences suggested that the subjects were tending to organize their calls in groups of five, with all five symbols frequently represented in five consecutive calls. To investigate this possibility, we considered each symbol position as a number response corresponding to its position (see Figure 1) and examined the distribution of *sums* of the positions called in sets of three, four, five, six, and seven calls ($k = 3, 4, 5, 6, 7$). The distributions of sums of each set of consecutive calls under uniformity were easily calculated and are shown, along with the observed distributions, in Figure 3. In all distributions the observed sums have a smaller variance than expected and a greater frequency of occurrence near the mean.

We were particularly interested in groups of five and found that our impression that the subjects tended to organize in groups of five was borne out by the highly significant excess of sums adding to 15 in the aggregate. Not only was this tendency found in the calls of all 82 subjects, but it was also remarkably stable from deck to deck within individual subjects.

The distinctive form of the composite observed distribution of sums of five-call sets (relative to sets of other sizes) was also, we may note, reflected in the distributions for individual subjects. That is, the spike in the middle of the distribution for $k = 5$ was not caused by relatively few subjects but represented a general tendency among subjects.

First and last calls

As we have said, each of the 82 subjects made 600 calls in 24 runs of 25 calls each. These runs were the experimental units. The first and last calls of each unit were tabulated, and the proportion of such calls for each position (based on the totality over subjects and decks) is given in Table 5. This table reveals less deviation from uniformity for the last calls ($\chi^2_4 = 10.68$) than for the first calls ($\chi^2_4 = 19.35$). The rank order of preference for the five positions corresponded with that for *all* singles in the case of first calls, but not in the case of last calls. In fact, the bulk of the contribution to $\chi^2_4 = 10.68$ was due to the deficiency of position-4 calls.

Table 6 gives some indication of the number of subjects with usual and unusual (some indeed unique) frequencies of response for their first and last calls in all 24 units. From this table it is clear that deviation from uniformity in making first calls was subject-related. For example, there were 5 subjects who made a first call of position 2 in more than half of their 24 units of 25 calls each. There were 7 subjects who never called

Table 5. First and last calls for each unit; proportions called for all decks and subjects

	Position					χ^2_4
	1	2	3	4	5	
First calls	.179	.230	.213	.198	.179	19.35
Rank order	5	1	2	3	4	
Last calls	.202	.215	.212	.173	.198	10.68
Rank order	3	1	2	5	4	
All singles	.183	.213	.208	.200	.186	
Rank order	5	1	2	3	4	

Note: Total number of first calls, 1,968; of last calls, 1,968 (82 subjects $\times$ 24 units).

position 5 as their first call, and 6 subjects who never called position 1 first. Among 82 subjects calling under uniformity, none would be expected to exhibit this kind of behavior in making calls. Table 6 also reveals that there was much less deviation from uniformity among subjects in their last-call behavior.

Table 6. First and last calls for each unit; analyzed by subjects

		Position				
		1	2	3	4	5
Number of subjects omitting the position	First call	6	2	2	4	7
	Last call	2	0	0	1	0
	Expected	0	0	0	0	0
Number of subjects calling this position at least half the time	First call	3	5	1	0	1
	Last call	0	0	0	0	0
	Expected	0	0	0	0	0
Number of subjects calling this position four, five or six times	First call	23	30	29	25	27
	Last call	39	38	39	38	30
	Expected	45	45	45	45	45

Note: Under the hypothesis of uniformity of position call, none of the 82 subjects would be expected to omit a given position on all 24 calls of a given unit, nor to begin 12 or more of the units at a given position.

Matching calls in runs

Examination of correspondence in sequence of calls both within and among subjects is intriguing, but the possible number of combinations

was overwhelming. To simplify the comparisons, we examined only the series of within-deck calls beginning with the first call in each 25-card unit; that is, we compared each subject's deck 1 calls only with all other deck 1 calls, his deck 2 calls with all other deck 2 calls and so on. Further, if the first call of two sets did *not* match, no further matching of those sets was considered. Even with these restrictions, there were $4^2 C_2^{82} = 52,136$ among-subjects paired comparisons and $82 C_2^4 = 492$ within-subject paired comparisons for each deck.

The expected (random) probability distributions of matching runs of length 0 through length 25, given 25 consecutive calls and five possible responses, were calculated (data available from the authors). The top part of Table 7 shows the observed and expected frequency of matched calls for these lengths, by deck, for the among-subject comparisons. It shows, for example, that in deck 1 (numerals, first sitting) there were five instances in which the first seven calls of a run were identical. Under random calling, this would have been expected only once. In deck 5 (open symbols, second sitting), there was one instance in which two individuals responded identically for the first ten calls of a 25-call unit. Table 7 also gives the *total* number of agreements observed and expected. One would expect 13,284 agreements for a given deck, and yet we observed substantially more than this for every deck.

The bottom section of Table 7 is just like the top, except that it is for within-subject paired comparisons. That is, for any given deck all possible paired comparisons were made among the four 25-call units of a given individual. These comparisons were then pooled for the 82 subjects making a total of $[(4)(3)/2] \times 82 = 492$ within-subject paired comparisons. In all cases, there was more agreement than one would expect due to chance. For within-subject comparisons, one would expect 123 agreements for a given deck, yet we observed a total of 1,340 agreements relative to the 738 (123×6) expected.

Although the analyses above do not indicate position calls or patterns of calls that are time- or deck-dependent, some of the particular call characteristics mentioned were examined in an effort to assess reliability in the test-retest context. The correlation between test-retest results for the 82 subjects were examined for (a) number of calls in extreme lateral positions (extreme-position calls); (b) number of groups of five consecutive calls in which there were no position calls, or symbols, repeated (no-repeat groups); (c) number of groups of five consecutive calls, the sum of positions of which equal 15; (d) number of like-membered doubles (11, 22, 33, 44, 55) called; and (e) number of doubles composed of unlike extreme-position calls (positions 1, 5, and 5, 1).

Table 7. Matching runs called for the various decks

Deck	Length of matching run																	Total agree- ments per deck
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Among-subject comparisons																		
1	42,008	7,968	2,272	596	174	80	26	5	4		2	1						15,650
2	42,642	7,822	1,875	527	146	92	15	10	5		1		1					14,419
3	42,583	7,787	1,906	568	176	83	21	7	3	1			1					14,642
4	42,434	7,797	2,086	536	170	80	22	5	3		2		1					14,880
5	42,458	7,834	1,924	577	147	142	36	6	9	2	1							15,069
6	42,588	7,861	1,901	473	144	124	30	7	5	1	2							14,576
Expected	42,508	8,502	1,700	340	68	14	3	1										13,284
Within-subject comparisons																		
1	384	69	17	14	1	4	1	2										189
2	370	68	29	10	3	6	4	1		1								239
3	388	53	19	14	4	10	2	1			1							229
4	390	56	24	9	2	2		2	4		1					2		238
5	391	63	14	5	4	10	3			1					1			214
6	377	67	22	12	1	11			1								1	231
Expected	394	79	16	3	1													123

Note: The data are for frequency, f , of matched sequences of length x ; the total numbers of agreements, $\sum f/x$, calculated by using the expected values given and multiplying by x , are slightly higher due to rounding error.

The correlation coefficient for each of these examples (Table 8) was highly significant ($p < 0.001$), but the differences among these correlation coefficients were not significant. This does not furnish evidence that some calling tendencies are more repeatable than others. On the other hand, the test-retest correlation coefficients calculated by deck (also Table 8) are further evidence of the consistency in calling behavior regardless of the type of symbol used.

Table 8. Test-retest correlation coefficients

	Extreme- position calls	No-repeat groups of five	Position sum = 15	Like- membered doubles	Unlike extreme- position calls
Over all card types	.74	.67	.63	.77	.67
Numerals only	.63	.50	.42	.53	.39
Open symbols only	.55	.53	.50	.73	.50
Closed symbols only	.55	.63	.56	.75	.60

DISCUSSION

The sections just above demonstrate a number of significant similarities in subjects' calls and sequences of calls. There were obviously far more differences between those calls than there were similarities. It is this very fact that makes for great difficulty in presenting the data, since differences may contribute as much to an understanding of the process of choice as do similarities. Each record appears to be an expression of that person's individuality and uniqueness. It differs markedly not only from all other subjects' records but also from what would be expected under uniformity.

In spite of the uniqueness of each protocol, few subjects exhibited markedly unusual records. For example, only one subject called significantly more extreme- than midvalue-position calls. However, there was considerably more variability in the degree of avoidance of like-membered doubles. This tendency toward individual variability was even more apparent in triples. Here again there was the paucity of like-membered calls, with much subject variation in the type of triples called, but each record showed more group conformity than deviance.

Though it is not practical to present the data idiographically, the individual protocols showed that strikingly many methodologies were used. Many of the subjects had attended courses in statistics in college

and medical school. Though they could not be considered students of probability, they had an acquaintance with its fundamentals. Yet, even though they knew that varying their strategy in calls did not increase the likelihood of their successfully identifying random numbers, each of them did vary his strategy. Most of them did not believe in extra-sensory perception and many did not even care to learn of their test performance. Yet, all showed immense 'nonrandom' variety in calls. Stereotyped, perfunctory repetition of numbers or strategies was infrequent.

It is not our purpose here to speculate on the motivation of our subjects¹ in order to explain their failure to call extreme positions or like-membered sequences as often as one expects under uniformity. Rather, we want to point up the fact that all subjects displayed characteristics that indeed did "have the quality and creases of well practiced rule-governed routines" (Bruner, 1966, p. 2). Yet, intertwined with these routines were unique subject variations that formed stable aspects of a given subject's test behavior.

Notes

Drs. Orr and Maxwell are in the Department of Psychiatry, and Dr. Federspiel is in the Department of Preventive Medicine (Biostatistics), at the Vanderbilt University School of Medicine, Nashville, Tennessee 37203. Received for publication March 1, 1971.

1. In a further study, data were gathered on the correlation between individuals' calling behavior and certain psychological and personality variables.

References

- Bakan, Paul. 1960. Response tendencies in attempts to generate random binary series. *American Journal of Psychology* 73:127-131.
- Bruner, J. S. 1966. *Toward a theory of instruction*. Boston: Harvard University, Belknap Press.
- Goodfellow, Louis D. 1938. A psychological interpretation of the results of the Zenith radio experiments in telepathy. *Journal of Experimental Psychology* 23:601-623.
- Goodfellow, Louis D. 1940. The human element in probability. *Journal of General Psychology* 23:201-205.
- Jenkins, W. O., and Cunningham, Leta M. 1949. The guessing-sequency hypothesis, the spread EFFECT, and number guessing habits. *Journal of Experimental Psychology* 39:158-168.
- Kendall, M. G., and Stuart, Alan. 1963. *The advanced theory of statistics*, Vol. 1. London: Charles Griffin and Co., Ltd.

- Nagel, Ernest. 1955. Principles of the theory of probability. In *International encyclopedia of unified science*, Vol. 1, part 2. Chicago: University of Chicago Press.
- Piaget, J., and Inhelder, B. 1951. *La genèse de l'idée de hasard chez l'enfant*. Paris: Presses Universitaires de France.
- Rath, Gustave J. 1966. Randomization by humans. *American Journal of Psychology* 79:97-105.
- Reichenbach, Hans. 1949. *The theory of probability*. Berkeley and Los Angeles: University of California Press.
- Rhine, J. B.; Pratt, J. D.; Stuart, C. E.; Smith, B. M.; and Greenwood, J. A. 1940. *Extrasensory perception after sixty years*. New York: Henry Holt.
- Ross, Bruce M. 1955. Randomizing of a binary series. *American Journal of Psychology* 68:136-138.
- Senders, Virginia L. 1953. Further analysis of response sequences in the setting of a psychophysical experiment. *American Journal of Psychology* 66:215-228.
- Yule, G. Udney, and Kendall, M. G. 1950. *An introduction to the theory of statistics*. London: Charles Griffin and Co., Ltd.

continued from page 80

- Reuchlin, Maurice. *Traite de psychologie appliquée*. Paris: Presses Universitaires de France, 1971. Pp. 240. Fr24.
- Richardson, Herbert W. *Nun, Witch, Playmate: The Americanization of Sex*. New York: Harper & Row, 1971. Pp. xii, 147. \$4.95.
- Rothman, David J. *The Discovery of the Asylum: Social Order and Disorder in the New Republic*. Boston: Little, Brown and Co., 1971. Pp. xx, 376. \$12.50.
- Rubin, Richard D.; Fensterheim, Herbert; Lazarus, Arnold A.; and Franks, Cyril M., eds. *Advances in Behavior Therapy*. New York: Academic Press, 1971. Pp. xv, 301. \$14.50.
- Sarnoff, Irving. *Testing Freudian Concepts: An Experimental Social Approach*. New York: Springer Publishing, 1971. Pp. xi, 276. \$8.50; paperback, \$5.95.
- Saul, Leon J. *Emotional Maturity: The Development and Dynamics of Personality and Its Disorders*. 3d ed. Philadelphia: J. B. Lippincott, 1971. Pp. xviii, 398. \$12.50.
- Schoonmaker, Alan N. *A Students' Survival Manual, or How To Get an Education Despite It All*. New York: Harper & Row, 1971. Pp. xvi, 373. Paperback, \$3.95.
- Schutz, William C. *Here Comes Everybody*. New York: Harper & Row, 1971. Pp. xviii, 295. \$6.95.
- Shaffer, Thomas L. *Death, Property, and Lawyers*. New York: Dunellen, 1971. Pp. xi, 292. \$7.95.
- Shaw, Lynette, and Sichel, Herbert S. *Accident Proneness: Research in the Occurrence, Causation, and Prevention of Road Accidents*. Oxford: Pergamon Press, 1971. Pp. xiv, 476. \$16.00.
- Shean, Glen D., ed. *Studies in Abnormal Behavior*. Chicago: Rand McNally, 1971. Pp. xi, 482. Paperback, \$6.95.
- Sheridan, Charles L. *Fundamentals of Experimental Psychology*. New York: Holt, Rinehart and Winston, 1971. Pp. xii, 402.
- Sherman, Julia A. *On the Psychology of Women: A Survey of Empirical Studies*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xi, 304. \$13.50.
- Shostrom, Everett, and Kavanagh, James. *Between Man and Woman: The Dynamics of Intersexual Relationships*. Los Angeles: Nash Publishing, 1972. Pp. 285. \$6.95.
- Shouksmith, George. *Intelligence, Creativity, and Cognitive Style*. New York: Wiley, 1971. Pp. 240. \$6.50.
- Singer, Jerome L., ed. *The Control of Aggression and Violence*. New York: Academic Press, 1971. Pp. xvii, 171. \$9.50.
- Singer, Laura J., and Buskin, Judith. *Sex Education on Film*. New York: Teachers College Press, 1971. Pp. vi, 170. Paperback, \$3.95.
- Slobin, Dan I., ed. *The Ontogenesis of Grammar*. New York: Academic Press, 1971. Pp. xi, 247.
- Stricker, George. *Kicking It*. New York: Pyramid Books, 1971. Pp. 159. Paperback, \$.95.
- Sutherland, John D., ed. *Towards Community Mental Health*. New York: Barnes & Noble, 1971. Pp. xii, 130. \$7.75.

continued on page 120

EVIDENCE FOR SHORT- AND LONG-TERM MEMORY IN MONKEYS

Douglas L. Medin
Rockefeller University

Performance by monkeys on a 16-choice indirect delayed-response task displayed an initial drop with time but approached an above-chance asymptote. A simple model assuming both short- and long-term memory fits the data closely. Under a variety of conditions affecting overall performance, estimates of the short-term component remain constant.

Reanalysis of a previously published study (Medin, 1969) suggests that memory in monkeys contains discrete short- and long-term components. In that study, nine experienced rhesus monkeys were tested on a 16-choice indirect delayed-response test. Each trial proceeded as follows: (a) the animal observed the brief illumination of one of 16 cells that were arrayed to form a 4-by-4 matrix through a clear Plexiglas screen; (b) a delay interval of 0, 1, 2, 5, 10, or 20 sec ensued; (c) the clear screen was raised and the monkey was allowed to respond to one of the 16 cells. If the monkey responded to the cell previously illuminated, it got a raisin reward and the light was turned on again. If it did not respond to the correct cell, no light was turned on nor any reward given. A more detailed methodology is outlined in the original paper.

The overall results showed that the proportions of responses correct were .937, .856, .740, .573, .482, and .437 for delays of 0, 1, 2, 5, 10, and 20 sec respectively. Chance performance would be 1/16 responses correct. Models assuming a unitary memory trace fading at some constant rate toward zero cannot predict these data accurately. In order to predict the efficient performance at the longer delay intervals, the predictions of these models must greatly overshoot observed performance at the shorter intervals.

A very simple model assuming both short- and long-term components in memory can adequately account for this result. The model assumes that if the monkey attends to the cell illumination, there is some probability that information sufficient to produce a correct response will enter long-term memory, from which it does not decay (Shiffrin and Atkinson, 1969). And if the monkey attends to the illumination, it is likely

that information sufficient to produce a correct response will enter short-term memory, from which it does decay (or is interfered with) at some constant rate with time. The conceptualization is consistent with a larger number of models, such as Estes' (1955) stimulus-fluctuation model. From these assumptions we can estimate the proportion of responses correct as a function of delay interval by Equation 1:

$$P(\text{correct}) = [LTM + (1-LTM)K^t]P(A), \quad [1]$$

where *LTM* represents the long-term component, *K* is the short-term decay parameter, *t* is the number of seconds of delay, and *P(A)* is the probability that the subject attended on a trial (this is estimated from performance on 0-sec delay intervals). In estimating *LTM* and *K* we adjust the scores for guessing.

Best-fitting values of *LTM* and *K* were estimated by a least-squares criterion and were found to be .43 and .78 respectively. With these values Equation 1 fits these data quite accurately, with the average absolute deviation of predicted from observed proportion correct being .01.

A more detailed analysis of the data reveals the further utility of this distinction between *LTM* and *K*. To begin with, performance in the 16-choice task improved substantially with practice. In terms of Equation 1 we can ask which of the two components was responsible for this improvement. A second type of analysis involves applying Equation 1 to the finding that all cell positions were not equally difficult: cells on the corners of the 4-by-4 matrix were easiest to remember, cells on the edge but not in the corners were less easy to remember, while cells in the middle four cells of the matrix were most difficult to remember.

Table 1 shows the best-fitting parameter values for *LTM* and *K* for each of these practice and cell-position conditions. Proportion correct at the 20-sec delay interval is also tabled to show the large performance differences these conditions produced. The average absolute deviation of predicted from observed values is included in the table, and in no case were the predicted and observed values significantly different.

The most striking result is that while the best-fitting estimates of the long-term memory component range from .20 to .68, estimates of the short-term decay parameter remain remarkably constant across conditions, varying only between .78 and .80. In terms of our two-process model, it appears that practice and cell-position (pattern) difficulty involve changes in long-term memory and do not affect short-term memory. The confinement of parameter changes to the long-term component is consistent with theories that attribute the major flexibility of memory to long-term memory processes (Shiffrin and Atkinson, 1969).

Table 1. Parameter estimates for some of the conditions from the Medin (1969) experiment; also shown is performance at the longest delay interval for each condition

Condition	Proportion correct at 20-sec delay	Long-term memory parameter (LTM)	Short-term memory parameter (K)	Predicted versus observed average absolute deviation
Practice				
First 5 days	.394	.37	.78	.015
Last 5 days	.492	.52	.80	.240
Cell position				
Center	.232	.20	.78	.028
Edge	.386	.36	.79	.006
Corner	.690	.68	.79	.010

In sum, the present analyses strengthen the hypothesis that it should be useful to distinguish between short- and long-term memory in monkeys. At the very least, a two-process model is needed to account for these data. As such, the data support the same distinction that physiological evidence suggests (see Kimble, 1965). In the present experiment, a simple two-process model of memory accurately described performance under a variety of conditions. Furthermore, improvement with practice and impairment with cell-position difficulty both involved changes in—and only in—the long-term component of this model.

Notes

This research was partially supported by USPHS Grant GM 16735. Received for publication February 10, 1971.

References

- Estes, W. K. 1955. Statistical theory of distributional phenomena in learning. *Psychological Review* 62:369–377.
- Kimble, D. P., ed. 1965. *The anatomy of memory*. Palo Alto, Calif.: Science and Behavior Books.
- Medin, D. L. 1969. Form perception and pattern reproduction by monkeys. *Journal of Comparative and Physiological Psychology* 68:412–419.
- Shiffrin, R. M., and Atkinson, R. C. 1969. Storage and retrieval processes in long-term memory. *Psychological Review* 76:179–193.

continued from page 116

- Timms, Duncan W. G. *The Urban Mosaic: Towards a Theory of Residential Differentiation*. Cambridge Geographical Studies, 2. New York: Cambridge University Press, 1971. Pp. viii, 277. \$11.50.
- Turney, Bill, and Robb, George. *Research in Education: An Introduction*. Hinsdale, Ill.: Dryden Press, 1971. Pp. xi, 320.
- Voneche, Jacques. *Les effets consecutifs figuraux*. Louvain: Vander, 1971. Pp. 226. \$4.80.
- Watts, W. David, Jr. *The Psychedelic Experience: A Sociological Study*. Beverly Hills, Calif.: Sage Publications, 1971. Pp. 256. \$8.95.
- Welkowitz, Joan; Ewen, Robert B.; and Cohen, Jacob. *Elementary Statistics for the Behavioral Science*. New York: Academic Press, 1971. Pp. xvi, 271. \$8.50.
- Wiggins, Jerry S.; Renner, Edward K.; Clore, Gerald L.; and Rose, Richard J. *The Psychology of Personality*. Reading, Mass.: Addison-Wesley, 1971. Pp. xii, 705. \$10.50.
- Willett, T. C. *Criminal on the Road: A Study of Serious Motoring Offences and Those Who Commit Them*. New York: Barnes & Noble, 1971. Pp. xv, 343. \$6.75; paperback, \$3.00.
- Wing, W. Cliff, Jr., and Wallach, Michael A. *College Admissions and the Psychology of Talent*. New York: Holt, Rinehart and Winston, 1971. Pp. ix, 165.
- Winick, Charles, and Kinsie, Paul M. *The Lively Commerce: Prostitution in the United States*. Chicago: Quadrangle Books, 1971. Pp. ix, 320. \$8.95.
- Wolfe, Dael. *The Uses of Talent*. Princeton, N.J.: Princeton University Press, 1971. Pp. x, 204. \$6.50.
- Wright, Derek. *The Psychology of Modern Behaviour*. Baltimore: Penguin Books, 1971. Pp. 288. Paperback, \$1.65.
- Yukl, Gary A., and Wexler, Kenneth N., eds. *Readings in Organizational and Industrial Psychology*. New York: Oxford University Press, 1971. Pp. xiii, 590.
- Zander, Alvin. *Motives and Goals in Groups*. New York: Academic Press, 1971. Pp. xiv, 212. \$11.00.

INPUT TESTING IN THE DETECTION OF MISPELLINGS

Donald G. MacKay

University of California at Los Angeles

Two types of misspellings (phonetically compatible, *WERK* for *WORK*; or phonetically incompatible, *WARK* for *WORK*) were tachistoscopically presented to 24 subjects. Telling them what word would be misspelled increased the probability of detecting phonetically incompatible misspellings, but not phonetically compatible ones. The findings present difficulties for passive models of perception; an active, input-testing model is more promising.

It may seem surprising that spelling errors should provide a testing ground for general issues in perception. But current data on detection of misspellings contradict the general assumption that perception results from direct or passive activation of analyzers for orthographic input. Rather, an active or mismatch process seems to be involved, whereby perception results from a comparison of the input (at various stages of coding) with a self-generated model of that input.

Consider these two types of misspellings: *phonetically compatible* and *phonetically incompatible*. Compatible misspellings can be pronounced the same as the original word (say, *VURSE* for *VERSE*) whereas the incompatible ones cannot (*VORSE* for *VERSE*). MacKay (1968) showed that incompatible misspellings were easier to detect in a sentence than compatible ones, a finding in accord with a mismatch model of perception (after Teuber, 1960). According to this model, the context of the sentence leads the subject to expect a particular word at the phonetic level, and detecting the misspellings depends in part on a mismatch between this expectation and the internal pronunciation of the actual input. Since the internal pronunciation of phonetically compatible misspellings matches that of the expected word, these misspellings pass undetected at the phonetic level. But the expectation fails to match the internal pronunciation of phonetically incompatible misspellings, which are therefore easily detected at the phonetic level according to the mismatch model.

The present study was a replication and extension of the 1968 MacKay study. To control for exposure duration, the phonetically compat-

ible and incompatible misspellings were presented tachistoscopically. The subject's expectations were controlled by instructing or not instructing him to expect a particular word. According to the mismatch model, phonetically incompatible misspellings should be easier to detect in the instructed condition than in the uninstructed condition. And phonetically compatible misspellings should be harder to detect than incompatible ones in the instructed condition but not in the uninstructed condition.

We also presented misspellings that form new words; for example, the misspelling of GREET as GREAT. Our question was whether subjects would identify GREAT more readily if they expected GREET than if they had no expectation at all. The mismatch model predicts that these 'new word' misspellings should behave like phonetically incompatible ones, being easy to detect in the instructed condition because of an internal mismatch at the pronunciation level and perhaps at the semantic level as well. We will see that the failure of this prediction leads to a theoretical reformulation of the mechanisms underlying the detection of misspellings.

METHOD

—Subjects and procedure—Twenty-four UCLA undergraduates received course credit for their participation in the 40-minute experiment. Each subject was warned that the words to be presented would sometimes be correctly spelled and sometimes not. He was to write down exactly what he saw, guessing at the spelling if necessary.

The materials consisted of 14 correctly spelled words (say, INCITE), 14 phonetically compatible misspellings (INSITE), and 14 phonetically incompatible misspellings (INBITE). The stimuli were typed in capital letters on 46 5-by-4-in. cards, which were shuffled thoroughly for each subject. The two types of misspelling were formed by changing exactly the same letter in one of the correctly spelled words. The misspellings altered neither word length nor letter height. The first letter was always correct but the position of the substituted letter varied from word to word. In addition, 5 new-word misspellings were formed by changing exactly the same letter in the original, correctly spelled words (say, INVITE for INCITE). The new-word misspellings slightly exceeded the original words in average Thorndike-Lorge frequency, but this is irrelevant here, since we were interested in comparing each word with itself in the 'set' and 'no-set' conditions described below.

In the *no-set* condition, a warning tone sounded 4 sec before the stimulus appeared. The *set* condition was identical except that the subjects were verbally told what word to expect in the 4 sec between the tone and stimulus onset. The order of the set and no-set conditions was counterbalanced across subjects. Half the stimuli were presented in the set condition and the remainder

in the no-set condition (counterbalanced across subjects so that each stimulus appeared equally often in both conditions). In both conditions all the materials were first exposed for 120 msec and then for 160 msec.

—Apparatus—A two-channel tachistoscope (Scientific Prototype 800F) presented the materials. One channel displayed a fixation point, and the other exposed the stimuli for either 120 or 160 msec. The words were 14 min (visual angle) in height and 30.5 min to 61 min in length. The luminance level of the stimuli was about 78 ftL as measured with a Spectra brightness spot meter. The tachistoscope was connected in series to three general purpose timers: one triggered the stimulus onset 4 sec after a warning tone, another fixed the duration of the warning tone (1 sec), and the third began the next trial by triggering the warning tone 12 sec after offset of the stimulus. An HP audio oscillator (201C) generated the 2,000-Hz warning tone.

RESULTS

Detection of misspellings

Compatible misspellings were no harder to detect than incompatible ones in the no-set condition. Here the probability of detecting compatible misspellings (averaged across subjects, words, and exposure durations) was .27; incompatible misspellings, .26. Nor did the set condition have a significant effect on the recognition of compatible misspellings (p at the .50 level, Friedman two-way analysis of variance on ranks). But the set condition did facilitate detection of incompatible misspellings. Significantly more incompatible misspellings were detected in the set than no-set condition ($p < .05$, same test). Similarly, incompatible misspellings were significantly easier to detect than compatible ones in the set condition ($p < .05$, same test).

We categorized incorrect responses as incorrect-word responses and as incorrect-set responses. An incorrect-set response was scored whenever the subject erroneously responded with the 'set' word, and an incorrect-word response whenever he erroneously responded with a correctly spelled English word. Table 1 shows the frequency of incorrect-set and -word responses for compatible and incompatible misspellings. As can be seen there, incorrect-word responses were no more frequent for compatible than incompatible misspellings in the no-set condition, but much more frequent in the set condition ($p < .05$, Friedman two-way analysis of variance on ranks). This difference was only partly due to the slightly higher frequency of incorrect-set responses for phonetically compatible misspellings. In the set condition (also shown in Table 1), incorrect-word responses remained about 8% higher for compatible than incompatible misspellings when the incorrect-set responses were subtracted

Table 1. The probability of incorrect-word responses and incorrect-set responses given an incorrect identification of phonetically compatible and incompatible misspellings

	Incorrect-word responses		Incorrect-set responses	
	Compatible misspellings	Incompatible misspellings	Compatible misspellings	Incompatible misspellings
Set condition	.71	.56	.40	.33
No-set condition	.54	.53	.10	.09

out. Thus compatible misspellings were mistaken for correctly spelled words more often than incompatible misspellings, but only in the set condition.

Detection of correctly spelled words

'Set' words (say, *HERD*) and new-word misspellings (say, *HARD* for *HERD*) made up the correctly spelled words. The set words were correctly identified more frequently in the set than the no-set condition ($p < .01$, Friedman two-way analysis of variance on ranks). But the opposite was true of new-word misspellings, which were correctly recognized significantly more often in the no-set than in the set condition ($p < .05$, same test).

Incorrect-set responses and incorrect-word responses were scored as before, with the results shown in Table 2. Incorrect-word responses were no more frequent for new-word misspellings than set words in the no-set condition, but were significantly more frequent for new-word misspellings in the set condition (84% versus 0%). This difference is only partly due to the high frequency of incorrect-set responses for new-word misspellings (also shown in Table 2). With these incorrect-set responses subtracted out, 50% of the erroneous responses to new-word misspellings were incorrect-word responses (as compared to 0% for the set words).

Table 2. The probability of incorrect-word responses given an incorrect response to two types of normally spelled words, along with the probability of incorrect-set responses for new-word misspellings

	Incorrect-word responses		Incorrect-set responses
	Set words	New-word misspellings	New-word misspellings
Set condition	.0	.84	.34
No-set condition	.40	.41	.08

Apparently the subjects realized the new-word misspellings were words, but were unsure *which* words.¹

DISCUSSION

Our findings present difficulties for passive models of perception. But the mismatch model developed in the introduction also fails to explain certain aspects of our data. According to this model, new-word misspellings should be easier to detect in the set than the no-set condition. Since the reverse was true, we were forced to develop a new model to explain our results. Our model can be considered an input-testing theory (after MacKay, 1967; Miller, Galanter, and Pribram, 1960), since perception is considered to follow a series of tests on the input.

Four possible input tests in the set condition of our experiment are shown in Table 3. Exposure duration was undoubtedly too brief to allow complete application of the first test (Is the input identical to the set word at the orthographic level?). And assuming that only one or two of the remaining tests can be applied on a probabilistic basis, this model seems capable of handling virtually all of our results.

Consider the possibility that a new-word misspelling has only passed test 3 (Is the input identical to the set word at the phonetic level?), with-

Table 3. The outcomes of four possible input tests in the set condition for the four types of stimuli

Input tests	Phonetically compatible misspellings	Phonetically incompatible misspellings	Set words	New-word misspellings
1. Is the input identical to the set word at the orthographic level?	No	No	Yes	No
2. Is the input a familiar configuration (is it a word) at the orthographic level?	No	No	Yes	Yes
3. Is the input identical to the set word at the phonetic level?	Yes	No	Yes	No
4. Is the input familiar (is it a word) at the phonetic level?	Yes	No	Yes	Yes

out time for further tests. Given no further information, the subject's best response is that the set word was presented, which accounts for both the high frequency of incorrect-set responses and the low probability of detecting new-word misspellings in the set condition. Note too that phonetically compatible misspellings pass tests 3 and 4, making the set word the best possible response if time permits no further tests. The input-testing model therefore explains the high frequency of incorrect-set responses for phonetically compatible misspellings.

The model also explains the ease of detecting phonetically incompatible misspellings, which fail all four input tests. The high frequency of incorrect-word responses for phonetically compatible misspellings and new-word misspellings can be handled in similar fashion. Both classes of stimuli pass test 4 (Is the input a word at the phonetic level?), forcing the subject to respond with a word if no further tests can be run. Finally, the model explains why erroneous responses for set words in the set condition were usually phonetically compatible. If the exposure duration permits all but test 1 (Is the input identical to the expected word at the orthographic level?), only two responses are feasible in this condition: one a possible incorrect-set response, and the other a phonetically compatible response. By chance alone, some of the responses will fall into this latter class, explaining the tendency to make errors that are phonetically compatible with the set word.

An input-testing model also captures the difficulty in detecting the absence of silent E's, as in HORDE, or silent K's, as in KNIGHT (Corcoran, 1966, 1967, 1968), since the missing element is absent both in the phonetic test and in the internal pronunciation of the misspelling. The model likewise predicts difficulty in detecting the addition of silent letters as in CORDE.

The findings of Rommetveit (1968) also fit an input-testing model. Rommetveit found that when SHAR is exposed to one eye along with SHAP to the other eye, the subjects reported SHARP, as if the input passed all the tests for perceiving SHARP. Similarly, Day (1967) found that when RODUCT is presented to one ear, along with PODUCT to the other, the subjects heard PRODUCT, as if the input passed all the tests for hearing PRODUCT.

In conclusion, further research on input-testing models is recommended to discover the order (if any) in which the tests are usually run; to determine whether different individuals apply different tests to the same input, leading to subject-specific misperceptions (such as Freud's misreading of *Hasdrubal* as *Hamilcar*; 1914); and to determine the nature of input testing in other modalities.

Notes

This investigation was supported by UCLA Grant 2428 and USPHS Grant 166668-01. The author thanks K. Achevski for running the experiment and analyzing the data, and M. Friedman for lending the experimental equipment. Revision received for publication April 13, 1971.

1. A curious response bias was noted for set words in the set condition. Here the subjects frequently created phonetically compatible misspellings of the set word. For example, they misspelled WORK as WURK, ADJOURN as ADJURN, and SURVEY as SERVEY. Moreover, these responses were inventions rather than carry-overs from the experimental misspellings WERK, ADJERN, and SURVAY that could have appeared earlier in the session. This response bias suggests that the subjects realized that the stimuli matched the expected word at the phonetic level but were unsure of its orthographic representation.

References

- Corcoran, D. W. J., and Weening, D. L. 1968. Acoustic factors in visual search. *Quarterly Journal of Experimental Psychology* 20:83-85.
- Corcoran, D. W. J. 1966. An acoustic factor in letter cancellation. *Nature* 210:658.
- Corcoran, D. W. J. 1967. Acoustic factors in proof reading. *Nature* 214:851.
- Day, R. 1967. Fusion in dichotic listening. Paper read to the Psychonomic Society, Chicago.
- Freud, S. 1914. *Psychopathology of everyday life*, trans. A. Brill. New York: Penguin.
- MacKay, D. G. 1968. Phonetic factors in the perception and recall of spelling errors. *Neuropsychologia* 6:321-325.
- MacKay, D. M. 1967. Ways of looking at perception. In *Models for the perception of speech and visual form*, ed. W. Wathen-Dunn. Cambridge, Mass.: The M.I.T. Press.
- Miller, G. A., Galanter, E., and Pribram, K. D. 1960. *Plans and the structure of behavior*. New York: Holt.
- Rommetveit, R. 1968. *Words, meanings, and messages*. New York: Academic Press.
- Teuber, H. L. 1960. Perception. In *Handbook of physiology*, ed. John Field, Vol. 3, pp. 1595-1668. Washington, D.C.: American Physiological Society.

ANNOUNCEMENT

The people of Asia need books and journals

UNESCO has designated 1972 as International Book Year. The Asia Foundation's Books for Asian Students program has set a goal for the distribution of one million books and journals for this Year. Your help will be greatly appreciated.

Books in the sciences, technology, social sciences and humanities dated 1960 or later in good condition, and professional and technical journals in runs of ten years or more are needed.

Shipments and questions may be addressed to

Books for Asian Students

451 Sixth Street

San Francisco, California 94103

Donations of books and journals are deductible under special provisions of the tax laws. The donor's tax adviser can provide guidance for valuation and deductibility.

THE DANCING ARABESQUE: AN UNUSUAL VISUAL EFFECT

Isaac Behar

U.S. Army Medical Research Laboratory, Fort Knox, Kentucky

An achromatizing lens moved back and forth laterally in front of the eye produces a dissociation of stimulus areas of different hue. This effect is the result of the lens having both positive and negative refractive power depending upon wavelength. Secondary visual effects are also described.

The axial chromatic aberration of the human eye, over the wavelength range 380 to 700 nm, is about 2.40 diopters (Bedford and Wyszecki, 1957). When 578 nm is used as the reference wavelength, the eye is relatively myopic for shorter wavelengths and relatively hyperopic for longer wavelengths. Thus, when visual studies are conducted of apparent intensity or resolution of small chromatic stimuli, it is necessary to take into consideration, or to correct, the axial chromatic aberration of the eye. One way to achieve the latter is with the use of an accessory lens in which the axial chromatic aberration complements that of the human eye. The design for such an 'achromatizing' lens is provided by Bedford and Wyszecki; the refractive power (as measured by an AO Lensometer using a B&L high-intensity monochromator) of a lens triplet made from this design is shown in Figure 1, along with the chromatic aberration values of Bedford and Wyszecki.

If one views an intricate, multicolored (arabesqueline) pattern through the achromatizing lens, and the lens, held close to the eye, is moved laterally back and forth, the components of the arabesque pattern dissociate and move (almost dance) relative to each other. This striking visual effect is not an illusion but is based on the fact that the lens is at the same time both a positive and a negative lens depending upon the wavelength. As shown in Figure 2, a simple positive lens moved in front of an aperture (say, the pupil of the eye) results in a change in the optical direction of a point, a change which is opposite to the direction of movement of the lens ('against movement'), whereas a negative lens produces movement in the same direction ('with movement') as that of the lens. Since the achromatizing lens is positive for wavelengths longer than 500 nm, and negative for shorter wavelengths, the lens produces

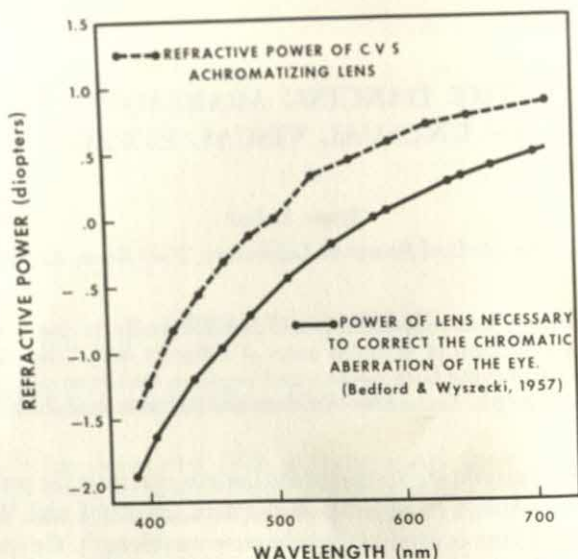


Fig. 1. The refractive power as a function of wavelength of the Center for Visual Science achromatizing lens and of a lens necessary to correct the axial chromatic aberration of the eye

'with' movement for blue and 'against' movement for red. The net effect is a literal dissociation of color patches with movement in opposite directions.

One consequence of this dissociated movement is an impression of depth. Since color patches of different wavelengths move in different directions and at different speeds relative to each other, the faster moving colors appear nearer, as one would expect from the operation of the motion-parallax cue. A second effect is the production of exaggerated borders when adjacent color patches move away from each other, and the minimization of borders when color patches move toward (or overlap) each other.

If an achromatizing lens is placed in front of the lens of an ordinary slide projector and an aperture of about 3 mm is moved laterally in front of the achromatizing lens, the dancing-arabesque effect may be demonstrated to a group. The effect may also be observed in the absence of an achromatizing lens by moving a small aperture in front of the eye. An especially effective pattern is provided by Figure 3-12 in Burnham, 1962 (which is taken from Plate XI of Evans, 1948). Here, the effect is considerably reduced because of the relatively small diam-

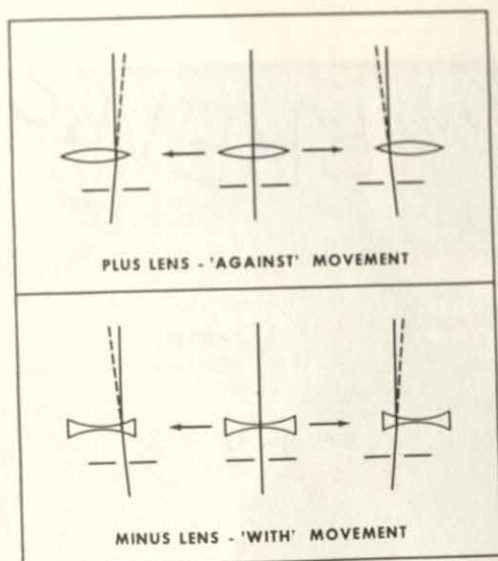


Fig. 2. Geometric diagram of change in optical direction of a point with movement of a positive and negative lens

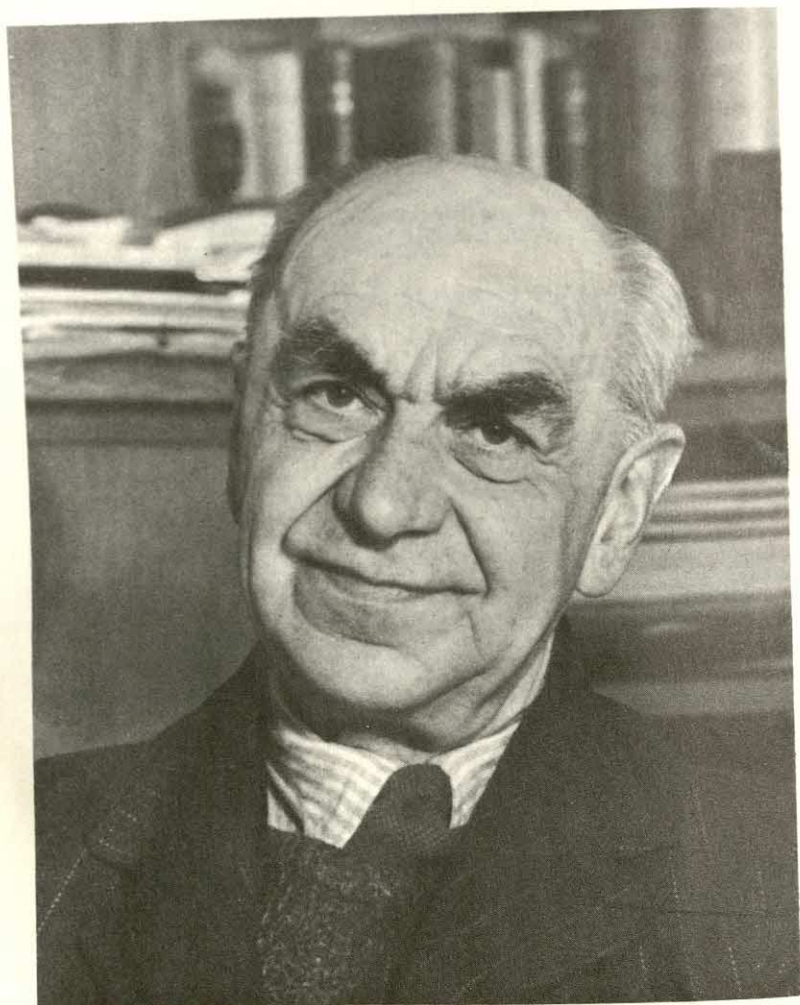
eter of the pupil, but still very evident. The use of a pinhole to observe the dissociated lateral movement resulting from the chromatic aberration of the eye has been described by Helmholtz (Southall, 1962).

Notes

The helpful comments of Dr. Robert M. Boynton of the Center for Visual Science are gratefully acknowledged. Received for publication March 12, 1971.

References

- Bedford, R. E., and Wyszecki, G. 1957. Axial chromatic aberration of the human eye. *Journal of the Optical Society of America* 47:564-565.
- Burnham, R. W., Hanes, R. M., and Bartleson, C. J. 1963. *Color: A guide to basic facts and concepts*. New York: Wiley.
- Evans, R. M. 1948. *An introduction to color*. New York: Wiley.
- Southall, J. P. C., ed. 1962. *Helmholtz's treatise on physiological optics*. New York: Dover.



J. G. Bartholomew

NOTES AND DISCUSSION

FREDERIC CHARLES BARTLETT: 1886-1969

Frederic Charles Bartlett was born on October 20, 1886, at Stow-on-the-Wold in Gloucestershire, a little country town where his father conducted a small but successful boot-and-shoe business. Bartlett has given some account of his early life there in his contribution to the third volume of Murchison's *Psychology in Autobiography*. This he greatly amplified in a short, unfinished, and largely autobiographical book on which he was engaged during the last few months of his life. Here is a fascinating picture of one little-appreciated element of the English social scene before World War I. Townships such as Stow, whose early importance and prosperity had long since diminished with the decline of the wool trade in southern England, and which had never been invaded by the Industrial Revolution, yet retained their integrity and enjoyed in their remoteness an independence and culture of their own. Religious nonconformity and political radicalism were sometimes significant features. Those, like Bartlett, who emerged from one of these townships might well carry a happy endowment of freedom from the social and intellectual constraints of the London-Oxford-Cambridge triangle, even though Oxford was but 30, and London 85, miles off.

But in Bartlett's case, accident reinforced whatever advantages an unconventional prelude to a Cambridge career may have conferred. A dangerous attack of pleurisy at the age of 14 put an end to family discussions whether he should be sent away to boarding school (the local grammar school having died, apparently of natural causes). He was left to educate himself, supported by his father's mental vigor and large assortment of books, by the stimulus, encouragement, and library of the local nonconformist minister, and by the freedom to immerse himself in the surrounding country life of farms, stone quarries, and village cricket—the latter to engender a lifelong passion and a curious intellectual stimulus.

Finally it was decided he should read for an external London University degree by taking the course offered by the University Correspondence College, whose headquarters were in Cambridge. On the casual suggestion of a friend of his father's, he took logic as one of his subjects. Of this choice he later wrote, "it seems odd that [an] apparently somewhat haphazard remark should have, indirectly, determined all my future, but so it turned out." Philosophy, including psychology, especially in the shape of James Ward's famous *British Encyclopaedia* article, engaged him, and his First Class Honours led to an invitation from the Cambridge Correspondence College to become Tutor in all philosophical subjects.

Cambridge

His work at the college was not so arduous as to prevent his taking the London M.A. and gaining special distinction in sociology and ethics, "though throughout this whole period his chief interest on the philosophical side continued to be Logic." In due course it became plain that he should venture on a fresh start as an undergraduate at Cambridge. Knowledge and admiration of the work of W. H. R. Rivers, who by then had moved from sensory physiology into anthropology, led him to Saint John's College, which already had a unique history in the encouragement of psychology through its Fellowship elections.

Bartlett's undergraduate career entailed much encounter with the prevailing variety of philosophical discussion, as well as some laboratory exercises in experimental psychology, largely conducted by Cyril Burt under the direction of C. S. Myers. Weight-lifting did not at the time appeal to him, though he valued the experimental approach whose discipline he did not fail to impose on his own students. Duly he achieved a First Class in his Tripos and though at the time his heart was set on anthropology, Rivers advised him that the opportunity, offered by Burt's departure, of taking charge of the experimental course would prepare him scientifically for a future career in the discipline of his choice. But the year was 1914 and circumstances outside the control of either of them dictated otherwise.

Many, notably Bertrand Russell and Maynard Keynes, have left impressions of those last luminous years in Cambridge preceding World War I. By some magic, the time needed for the intensive intellectual work on which the outstanding distinction of the University depended still seemed to allow ample leisure for games, country rambles, and the enlargement of personal relationships. For Bartlett, walks with Norbert Wiener (then studying with Bertrand Russell) were the occasion for much fruitful wrangling, and he tells us how Wiener threw out a suggestion that gave rise to Bartlett's most valuable method in the study of remembering—the method of serial reproduction.

World War I

But the war rudely shattered this engaging and fruitful world. Myers, Rivers, McDougall, and William Brown found their way into military psychiatry. Bartlett, without medical qualification and debarred by health from ordinary enlistment, remained in Cambridge as Myers' deputy in charge of the Psychological Laboratory. (In 1915, when the pressure of work on 'shell shock' cases became heavy, Bartlett tried his hand in this field but seems to have found incomplete satisfaction among patients who were constantly coming and going.) In the laboratory Mary Smith, his wife to be, was already at work, and the two collaborated on problems of listening to sounds of weak intensity connected with the operation of antisubmarine hydrophone equipment. This they continued together immediately after the war, and the two lengthy papers that resulted are of deep interest, partly because of the extended use of introspective methods, never again employed by Bartlett in a central role, and partly because of the light they throw on the earlier development of his views about feelings, attitudes, and orientations.

During the war, too, Bartlett wrote up in thesis form the experimental material later published in *Remembering*. This extended thesis, which also stated his theoretical position in those days, especially in relation to German

and French workers of the early twentieth century, gained him a Fellowship at Saint John's in 1917. The present-day reader will be surprised by two of its features. The first is the extent to which the empirical data presented in *Remembering* had already been collected and analyzed 15 years before the publication of the book. The second is the magnitude of the later change in his theoretical thinking about memory, chiefly occasioned by his close acquaintance with Sir Henry Head in the early 1920s.

The Cambridge laboratory

As things turned out, Rivers, after a brief but inspiring return to Cambridge, died in 1922. In the same year Myers withdrew from academic life to found the National Institute of Industrial Psychology in London. Bartlett was left in sole charge of the Psychological Laboratory, and the 1920s were spent in building a well-balanced department with strong emphasis on experimental work. Physically, for the needs of that time, the accommodation in a specially built wing of the Physiological Laboratory was excellent—the legacy of Myers' prewar efforts and personal generosity. Three Lectureships, in psychopathology, experimental psychology, and animal behavior, were established; and on the research side, support was given from the start by the Industrial Health Research Board and the Medical Research Council.

With all this in train and a constant flow of able students, Bartlett could give part of his attention to some psychological aspects of his earlier love—social anthropology. The possibility of fieldwork had by now clearly to be abandoned, but the views he had derived from his earlier studies of remembering were aptly brought to bear on some of the problems of culture contact and of the transmission and transformation of material (such as myths and folklore) in those social interactions. This found expression in his *Psychology and Primitive Culture*, published in 1923, and in a series of papers in the next few years. One of these formed his contribution to the Seventh International Congress of Psychology at Oxford in 1923.

That Congress was also the occasion for Bartlett, hitherto untraveled, to make numerous personal contacts and a number of lasting friendships with contemporaries in other lands. Here he first met Albert Michotte, Walter Miles, Edwin G. Boring, Herbert Langfeld, Wolfgang Köhler, and Kurt Koffka, to name but a few with whom his contact later extended to cover their respective junior colleagues. Henry Head too, an intimate friend of Rivers, attended this Congress, and the relationship between him and Bartlett during the first half of the decade formed, as we shall see, a highly significant factor in the development of Bartlett's thinking. Contact was close, for at that time Head was engaged in writing his *Aphasia and Kindred Disorders of Speech* (1926) and was in the habit of inviting Bartlett (by telegram) to come and discuss portions of the draft.

By 1930 psychology, in its experimental guise, was firmly established, as the University recognized by instituting the Chair in 1931. The laboratory, though small by American standards, was full of varied life. Staff and students, undergraduate and graduate, crowded Bartlett's room for the twice-weekly discussion classes that covered a wide range of topics, and echoes of the proceedings rebounded in the corridors and rooms of the laboratory—not to mention the 'local'—betweenwhiles. Research was equally varied; and although workers were free to choose their own interests, these reflected a

pattern representative of Bartlett's own. Auditory fatigue, memory, learning, thinking, visual perception, the effect of environmental distraction on skills, and animal behavior were all being actively investigated, Bartlett himself always ready to act as a (distinctly human) guinea pig.

The year 1932 saw the publication of *Remembering*, perhaps more of a turning point in Bartlett's own development than may have been apparent at the time. On the empirical side, the book made public the extensive findings previously expounded in his Fellowship thesis, which definitively established the predominantly reconstructive activities involved in recall, and it demonstrated the process of rationalization and conventionalization that largely derive from the individual's social and cultural setting. But the book differs in two very marked respects from the thesis. First, in place of the rather tentative interpretations based on his extensive study of previous French and German contributions, we find the theory of 'schemata,' which is the direct, if heavily and somewhat obscurely transformed, outcome of theoretical views formulated by Henry Head (not in *Aphasia*, where he never mentions the term, but in his *Studies in Neurology* of 1920). A second novel and outstanding feature of *Remembering* is the long and interesting discussion of the social and cultural aspects of memory, a reflection of Bartlett's continued preoccupation with earlier anthropological enthusiasms.

Students and others wrestled earnestly with the difficulties, and with what seemed to them the fascinating and novel potentialities, of the notion of schema, but on the whole their efforts to clarify the essential elements of the theory so as to make it applicable to further use in empirical investigation and experiment were unsuccessful, and Bartlett himself carried the matter no further. Mere criticism of it presents little difficulty; yet in the minds of at least some later workers, the idea has continued to present features that are lacking in other formulations of the role of past experience in determining present behavior. If the theory of schemata was in the last resort a failure, it was certainly one of real quality in imagination and insight.

With *Remembering* off his hands, Bartlett himself was concerned to effect a synthesis of anthropology and psychology, and so place social psychology on a fresh and firm basis. He took the leading part in organizing a small group of sociologists, psychologists, and anthropologists, a group which met twice yearly between 1935 and 1938 to try to hammer out some common principles of research method, rather than of doctrine. The meetings, often held in healthily remote places, produced excellent discussion between people who otherwise would not have come into close contact. The aim was to produce a book which—as its format, if nothing else, indicates—was to resemble a modern *Notes and Queries for Anthropologists*. But the outcome could not match that famous model in coherence and consistency. In any case, it was published in 1939, and the war swallowed it up along with much else.

World War II

In the second half of the 1930s two factors set Bartlett on the path that was in a sense final, since in one way and another, and especially at the organizational level, he pursued it for the rest of his life. In the first place, the eventual outbreak of war became more and more evident, and he was drawn into official discussions, especially of the psychological problems revealed by the expansion of the Royal Air Force.

At about the same time, in 1936, Kenneth Craik joined the junior research staff of the laboratory. This most unusual young man, himself endowed with extraordinary gifts both theoretical and practical, stimulated Bartlett's thinking along lines hitherto nascent—especially the theory of bodily skills, though this must have long preoccupied him in connection both with the theory of schemata and with games such as cricket and tennis. Manifestly, in a war that (at any rate, at the outset) would depend heavily on the use of aircraft, factors affecting the skills involved in flying would be of outstanding consequence. As a member of the Air Ministry's Flying Personnel Research Committee, Bartlett took a leading part in explicating the problems involved and in initiating the appropriate laboratory research.

In this, the close association with Kenneth Craik was indispensable. For all Bartlett's powers of perceiving essentials and exercising insight, he was neither by training nor inclination the man to realize in concrete and precise form either the theoretical conceptions involved in, or the apparatus required to carry out, the necessary experiments on factors affecting the behavior of pilots and the effects of fatigue. Craik was the perfect complement, for he could not only brilliantly sharpen and extend Bartlett's thinking but was also an engineering craftsman of quite unusual caliber. His original, discerning, and versatile powers were called on in the ever-greater variety of problems cropping up throughout the war, and other able workers were gathered into the laboratory to handle the constantly increasing volume of work. In 1944 the Medical Research Council established this group as the Applied Psychology Research Unit under Craik's direction. The latter's death in a road accident two days before the end of the war in Europe was a tragic loss to his science and his friends. For Bartlett it must have been a blow that few, if any, could even begin to appreciate.

The work of the Applied Psychology Unit continued and expanded, for the types of problems that arose in the stresses of war were found no less obtrusive in the new, more and more technologically oriented, postwar world. Bartlett himself devoted some considerable attention to formulating his views on the character of skills, though he never put together the various aspects of his thinking in any single monograph. Nevertheless, from a number of papers and lectures, notably his Ferrier Lecture in 1941 and the Oliver-Sharpey Lectures in 1947, there emerges the picture of 'skill' as an activity that, though based on a distillation of past experience and practice, remains flexibly adapted to moment-to-moment circumstances, involves anticipation, and depends on a grading of response in accordance with the external situation to be contended with. Above all, it involves correct timing of the various components that go to make up the skill—timing is vital. Breakdown in the face of fatigue (or other less transitory conditions such as neurosis) manifests itself in mistiming on the one hand and faulty grading of response on the other.

Thinking

Now, however, Bartlett's interests moved afresh, though once again not without connection with previous phases of his development. This time it was thinking that attracted his attention. The origins of this interest may well have been threefold.

First, so far as the empirical aspect was concerned, he had long entertained the idea that the methods reported in *Remembering* could be adapted

and extended to the study of certain kinds of thought processes. Thus, in place of a verbal passage complete in itself for the subject to recall, there could be a description of an open-ended situation, the subject's task now being to complete the story in the way he considered most plausible. This could reveal the factors entering into very common everyday situations in which the individual, often unawares, foresees the outcome of something as yet unrealized. Secondly, and connected with this, it is pretty evident that the notion of the schema as an active organization of past experience calls for extension of its application to constructive and predictive mental processes. Third, and more remotely, his old fascination with what goes on in the skills involved in such games as tennis and cricket became assimilated to his experience of constructive thinking, an idea which even the unsporting may find stimulating and suggestive.

For Bartlett all this led in empirical terms to the view that, as a practical activity, thinking involves the completion of some previously incomplete state of affairs—whether by interpolation or extrapolation—and this approach he explored experimentally. The time he needed to work his way through this potentially rich field on a scale comparable to his study of remembering was, however, lacking. His book *Thinking: An Experimental and Social Study*, which appeared in 1958, is less remarkable as a scientific contribution than as a precious document which, apart from some interesting material, allows us a closer view of his personal attitudes and thinking than do any of his previous works.

Later years

In 1952 Bartlett retired from the Chair of Experimental Psychology and the Directorship of the Psychological Laboratory but did not relinquish his connection with the Applied Psychology Unit, now independent. For a number of years, too, he remained active in the world of committees, and his advice was constantly sought and freely given in many quarters, official and unofficial. A proportion of his pupils, unusually high for any departmental head, by now themselves occupied positions of responsibility, and many of them turned to him for advice, help, or refreshment. His outstanding gift for seeing unexpected but significant aspects of a problem or difficulty long remained with him.

So unusual and remarkable was Bartlett's distinction, it seems almost superfluous to record the public honors that came his way. He received honorary doctorates from the universities of Athens, Princeton, Louvain, London, Edinburgh, Oxford, and Padua. In America, the National Academy of Sciences and the Academy of Arts and Societies honored him; but it was perhaps his membership in the American Philosophical Society, with its long and distinguished history, that gave him the greatest pride and pleasure. Numerous national psychological societies conferred honorary membership on him. He became a Fellow of the Royal Society in 1932, received the Companionship of the British Empire in 1941, and the Baly and Huxley medals of the Royal Society in 1943. Knighted in 1948, in 1952 he was awarded that outstanding scientific distinction, the Royal Medal.

Though somewhat, and increasingly, troubled by deafness, he was blessed with good health and stamina, and when he died on September 30, 1969, at the age of 82, his last illness had been mercifully brief.

Personal and intellectual frame

Bartlett's presence was so immediately compelling and impressive that he wielded influence over territory far outside his own academic and scientific pastures; though few, perhaps, of the many people he thus encountered ever attempted any analysis of his mental and personal qualities. To those in closer touch, he presented what sometimes seemed a constantly changing enigma. Certainly one of his most marked characteristics was a complete disregard of the trivial and a neglect of such traditional trappings of academic or social intercourse as served no useful purpose. In discussion he was by turns disconcertingly direct and—sometimes obscurely—oblique. Many discovered the need for subsequent reflection before they could grasp fully the implications of what he had said. Much the same was true of what he wrote, the intellectual texture of which always repaid further perusal and consideration.

Rather similar double aspects may be found in the development of his thinking. An early element, as we have seen, was formal logic, but this became progressively less an overt feature of his style, although he never dispensed with logical structure. In the same way, under the influence of Rivers, he early embraced the formalities of laboratory psychology yet soon explicitly rebelled against their practice, while retaining respect for them where their application was appropriate. Inherent personal tendencies apart, these traits may have derived from his youthful admiration for the work of James Ward, who combined a rigor of intellectual analysis with a liberal and beneficial importation into psychology of broad evolutionary principles.

An abiding and fascinating characteristic of Bartlett's mind was that for him a theme had no single exclusive significance. Often, it seemed, his thinking was enriched by a duality of reference, sometimes remote and surprising. So, the changes undergone by the matter of experience as it remains stored away in memory and is subsequently recalled or otherwise used seem akin in his thought to the fate of cultural material orally transmitted from generation to generation or between interacting social groups. Again, the brilliant stroke at cricket, with its use of all the organized past experience of the batsman; the commitment to a decision at a point from which there is no return; and the coordination in time of several contributory processes—these seem to merge in his mind with that type of thinking in which, explicit logic absent, a subject's adventurous flash of insight is in fact the product of a combination of appropriate selection of features of a problem and of elements from the organized mass of accumulated experience. Sallies such as these were neither mere discursiveness nor consciously contrived analogies. They lay at the root of Bartlett's unusual way of thinking and proved their value by enabling him to stand outside the more compulsively organized logic of others and see the issue from a fresh and often revealing viewpoint.

As a person Bartlett could at times appear austere, impassive—even withdrawn. At other times, and more frequently as he got older, a robust and striking boyish gaiety would seize him. No one who once heard his laugh, sudden and often unexpected in its setting, or saw the naughty gleam of friendly mischief in his eye, would ever lose the memory of it. His concern for others, especially his juniors, was by no means always apparent to them, but those in whose interests he worked behind the scenes sooner or later became aware of its extent and depth. He was much on his guard against grati-

tude that might embarrass him or compromise his subsequent freedom of action.

For a man so involved in such a variety of duties and obligations, he was astonishingly accessible and, until World War II at least, disdained the protection of a secretary. Indeed, he did not have a secretary at all, and himself typed all that needed typing on an increasingly battered machine. A visit to his office to discuss any problem was a spontaneous, informal affair in which he would often appear to talk round the point at issue and sometimes indeed give the impression of failing to appreciate it. Yet one left with a curious impression that the matter had somehow been clarified, though time was often needed to perceive exactly how. Direct personal advice he gave only very rarely.

Bartlett's preeminent part in the development of psychology in Britain, and to some extent in other countries, was the outcome of a rich and in some respects paradoxical personal and intellectual makeup. All in the subject area admired his own contributions, especially in their empirical aspect. But by no means did all see eye to eye with him as to the lines along which the science ought to develop, or what its proper content ought to be. As he advanced in age, Bartlett's convictions about the character and role of psychology progressively took on a more and more determinate form.

From early in his career he had felt that it was important for psychology to emerge from an aseptic ivory tower and meet the challenges of the world around. To a very large extent, World War II and its aftermath afforded him this opportunity and he seized it eagerly—if along lines some might regard as eventually restrictive and as hardly covering the range of the possible. In the last year of his life so strongly does this conviction seem to have seized him that in the small, unfinished book already mentioned, in which he undertakes to relate the history of psychology in Cambridge, the reader may be startled to find no reference to anything whatsoever between 1920 and 1939. Outstanding and distinguished as was his contribution during and after World War II, this omission does the gravest injustice to much that was most characteristic and fruitful in his own work. It may, indeed, be to the personal contributions he made during this middle period that workers of the future will turn for refreshment and stimulus, rather than to the well-established and, in some of their remoter branches, occasionally arid results of his later activities.

Throughout, his true achievement was in a major sense immediate and personal. This is true not only of his own work but of his impact on his students and successors. Largely, they were not engulfed but gained their own individuality through healthy reaction with him, an outcome well in keeping with the firm valuation of freedom so essentially engrained in his nature.

R. C. Oldfield, *University of Edinburgh*

Note: Professor Oldfield expresses deepest gratitude to Lady Bartlett for so freely making available much vital material on which parts of this account are based, and for going to much trouble to respond to questions to which she alone knew the answers.

BOOK REVIEWS

Julian Hochberg
Columbia University

Visual Perception

By Tom N. Cornsweet. New York: Academic Press, 1970. Pp. xiii, 475. \$15.00.

Principles of Perception

By S. Howard Bartley. 2d ed. New York: Harper & Row, 1969. Pp. xx, 518. \$11.50.

Visual Perception of Form

By Leonard Zusne. New York: Academic Press, 1970. Pp. xi, 547. \$19.50.

Visual Pattern Recognition

By Peter C. Dodwell. New York: Holt, Rinehart and Winston, 1970. Pp. x, 276. \$10.50.

Psychophysical Analysis of Visual Space

By John C. Baird. New York: Pergamon Press, 1970. Pp. viii, 321. \$10.50.

These five books are a good cross section of what is happening in the field of visual perception today. Anybody familiar with the history of American psychology in this century will recognize that they could not have appeared as recently as 25 years ago.

It was then, from the mid-1940s on, that we began to rid ourselves of slavish obeisance to naive positivism. Only a few stalwarts, like S. H. Bartley and James Gibson, together with the newly arrived emigrants from Germany, had kept the study of perception alive during the decade before World War II. Then the very real contributions of sensory psychology to the war effort became apparent. Also, there was a growing tolerance for epistemological ambiguity. All of this led to renewed efforts to bring sensory psychology into the framework of current scientific prejudice. Hence, some workers began to think of reports of experience as being nothing more than overt responses. Others found more sophisticated support—in philosophical writings such as those by Gilbert Ryle, or perhaps in Hebb's theory, which, if one did not look too closely, seemed to fit with behaviorist attitudes.

The mere existence of these five books testifies to the fact that psychology has settled down to widespread and serious study of sense perception. This has happened despite the fact that the old philosophical problems which led to the exclusion of such studies in the past are not yet solved.

For some time now, there has been a need for a really good textbook in this newly revived field of perceptual research. I am happy to report that in *Visual Perception*, Tom N. Cornsweet has written the first 16 chapters of what may well be the best textbook ever done on perception. The book is nicely printed, beautifully illustrated, and in general, extraordinarily lucid in style. The only important thing wrong with it is that it is incomplete. It deals with the visual stimulus and with what we used to call visual sensation. It does not even get into binocular vision. But what it does it does so very well that from now on I shall require all of my students to read it.

Those of us who remember the dry-as-dust treatment of visual psychophysics that drove so many good people into less worthy fields of psychology need fear no longer. Cornsweet makes it clear and fascinating.

The book is based on a simple and expert treatment of the physical stimulus. Cornsweet avoids the complexities of the usual account of light as a stimulus. Those of us who are used to the convoluted and distracting discussions of footlamberts, millilamberts, radiometric units, and so on will be pleasantly surprised to see that they do not enter Cornsweet's discussion at all. He cuts through all of this and presents only the essential fact that light is energy and that the only effective form of this energy is what enters the eye. As energy it has magnitude or intensity. This is all that the student needs to understand many basic facts of visual sensitivity. Cornsweet describes this light energy as quanta that travel in straight lines from a source. The rectilinear propagation of these packets of energy allows the use of what amounts to a simple ray model of light. He shows how the rays travel through diverse media, and a student who is paying attention will discover how to trace the rays through optical systems, including the human eye.

Using the classical experiment of Hecht, Schlaer, and Pirenne as an example, Cornsweet shows graphically how the detection of light is related to the absorption of quanta by the photopigments. Factors that affect this probability of absorption include the size of the source, its intensity, the time of its presentation, and its spectral composition. Fluctuations in sensitivity at threshold levels and at suprathreshold levels are related to the essentially statistical nature of the absorption of quanta.

From a didactic standpoint, the chapters on color vision range from excellent to difficult. Cornsweet prefaces this section with a nice discussion of the changes in the visual pigments attributable to absorption of light. In this connection he describes the central experiments of Wald, Rushton, and Campbell. A typical device is his distillation of the different versions of Rushton's densitometric apparatus into a single ideal system that gets the basic idea across. But Cornsweet is no uncritical relator of scientific tales. Despite the fact that this book is designed for the beginner, he points out the uncertainties in the conclusions that one might want to draw from the densitometric data.

The book then turns to a discussion of wavelength discrimination and relates this in a sophisticated way to the facts of color naming. Beginning with monochromacy and dichromacy, Cornsweet shows how multiple pigments lead to wavelength discriminations and can account for the facts of color mixture. The notion of a wavelength-mixture space is made clear with many diagrams and examples. A charming section describes how a monochromat

may be transformed into a dichromat with a suitable pair of spectacles, and how these same spectacles can convert a normal trichromat into a tetrachromat.

Unfortunately, the chapter on the perception of color is much more difficult to read than the chapters on wavelength discrimination and retinal color systems. The chapter on the perception of color suffers from several defects. First, Cornsweet's discussion about how to maneuver within the tridimensional wavelength-mixture space is very unclear. The perspective drawings of this space are difficult to interpret. The author of a book on perception should have been aware that a little line within a perspective diagram may well appear to be flat on the page and not to be oriented in depth unless adequate cues are provided. Secondly, a very long footnote comparing the CIE diagram with the system contained in the text is of no help whatsoever to a beginner. One gets the impression that the footnote is there solely to show that the author knows about the CIE system. Although this particular chapter could have benefited from considerable editing to eliminate some of the complexities for the beginning student, this reviewer looked in vain for some account of binocular color mixture. This phenomenon was once of considerable importance for color theory. It should at least have been mentioned in terms of its bearing on the trichromaticity theory of color mixture.

The three chapters on spatial interaction in the visual system are a beautiful piece of work. They include the usual discussions of lateral inhibition and contrast phenomena. But this is related to the exciting developments in the spatial-frequency domain. This latter topic is one of those examples of how developments in the physical sciences have had an almost immediate and direct impact upon psychology. The modulation-transfer function, which is right out of the field of electrooptics, is made perfectly clear. There is a nice elementary treatment of how optical systems may be considered to be filters capable of excluding various Fourier components.

Cornsweet also provides an extremely useful account of the implications of assuming that the visual system responds linearly and of how these implications may be erroneous. Typical responses of nonlinear systems, say, half-wave rectification and logarithmic transformations, are used to illustrate the effects of nonlinearity in the visual system. This is particularly important because so many psychologists seem to be unaware of the essentially nonlinear properties of the organism. Though this is a very sophisticated discussion, there is no reason why a student with a good command of the English language cannot understand it. Cornsweet uses the concept of nonlinearity in accounting for departures from brightness constancy.

The chapter on flicker perception is quite easy to read after the ones on spatial interaction. The temporal modulation function is a direct analogue in the time domain of the modulation-transfer function in the space domain. Again, the importance of nonlinearity is noted and its implications are made clear.

Now that I have said all of these nice things about Cornsweet's book, I feel that I am allowed to discuss some of its negative characteristics. First, the book is mistitled. It should have been called *Visual Sensation and the Visual Stimulus*. There is one small chapter on higher-order processes, but I think

that it is naive. Cornsweet's preoccupation is clearly with the retina. In his system, the brain does very little for perception. There is no treatment whatsoever of the perception of space, movement, or form. These major areas cannot be left out of a book ostensibly dealing with perception.

Psychologists are terrible name droppers. When mentioning the inverse-squares law, they are likely to cite Newton and Hooke and appropriate seventeenth-century dates. It is obvious, therefore, that there are certain advantages to minimizing the number of references, particularly in an introductory book. Cornsweet, however, goes to extreme lengths to avoid them. He has, for example, left out all reference to Kuffler, whose work is featured very prominently in some of the chapters. There are many other worthy works in the fields of brightness perception and flicker perception that could have been cited but were not. I wonder to what extent this reflects an unconscious acceptance of the notion that one must be a member of a particular clique oriented around the Berkeley-Cambridge axis in order to be citeable.

S. Howard Bartley's book *Principles of Perception* attempts to cover much more ground than does Cornsweet's. I am sorry to say that it does not do this very well at all. The book contains 17 chapters that range from visual acuity, through brightness, color, and space perception, to audition, the skin senses, taste, and smell, to perceptual learning and social perception. In general I found the book a fairly dry and uncritical reportage of miscellaneous topics in perception.

Bartley defines perception as "the immediate discriminatory response of the organism to energy-activating sense organs" (p. 11). Bartley recognizes that the term 'immediate' needs definition, but I found his attempts to do so unconvincing. Imagine a bright spot appearing at position A. It is then turned off and another bright spot is made to appear at position B. If the conditions are right, an observer may say that he saw an object that started at A and moved to B. You could then ask him if it passed through position C on its way to B. He is likely to say yes, and what is more, he will report that it passed through C before it got to B. Yet we know that the light must be flashed at B before it can be seen to move through C. Hence the whole episode of seeing the light move from A through C to B must be reconstructed after the fact. There is no immediacy here of the kind Bartley describes. The whole notion that perception involves considerable prior serial and parallel processing has had little impact on this book.

In most of the chapters, there are no unifying discussions to tie various facts together. When such discussions do appear, as in the chapter on the development of perception, they are bland and too general to be very meaningful. Thus, the chapter in question ends with the comment that perception is changeable. No feeling is given for the vital and often bitter discussions on the validity of many of the experimental conclusions cited in this chapter. We might cite the important controversy between Weinstein and Held on the role of voluntary movements in perceptual adaptation in this connection. Also, the work of Harris, who disputes the idea that visual adaptation *per se* actually occurs, is ignored. Similarly, Rock's related theoretical ideas are left out. I suppose that Bartley must suppress this evidence if he is to maintain his view that visual perception is changeable.

There are several other surprising omissions in the book. The important and classical work on contour inhibition by Fry and Bartley is reported, for example. But apart from a brief account of the findings with the *Limulus* eye, there is no mention of lateral inhibition, nor is there any attempt to relate it to the contour-inhibition phenomenon. The chapter on color vision also suffers from important deficiencies. There is an entirely incomplete account of the color theories, which are just barely described. How is a student supposed to relate diverse experimental findings to each other in the absence of theoretical concepts? Similarly, the color solid is described, but no feeling whatsoever is given of how it was conceived or on what facts it was based.

In this chapter on the spatial constancies, Bartley has got his facts wrong about the moon illusion. It was Boring and Taylor who did not get a monocular effect, Rock and I did. Also, we showed an influence of degree of cloudiness, which Bartley says does not exist. When dealing with the constancies, Bartley leaves out the important kinetic cues to depth such as movement gradients, motion parallax, and the KDE.

The chapter on space perception is largely atheoretical. There is a nice explanation of disparity but nothing at all about theories of stereopsis—undoubtedly one of the more lively topics in today's psychology. Facts are given about accommodation and convergence, but some of the more interesting phenomena, such as the roles of astigmatism and chromatic error in guiding accommodation, are left out. The chapter does stress the importance of nonvisual information in space perception. This, in my opinion, is an advance over some of the older treatments. The material on brightness constancy is entirely inadequate. There is no mention of the work of Hess and Pretori, Wallach, Heineman, Hurvich and Jameson, or even Gelb.

In discussing form perception, Bartley does repeat the useful point that the stimuli in this field may actually be responses. Some of the work of Attneave, Arnoult, and Boynton is cited. But, again, no theory. The illusions are discussed briefly in this chapter, but we learn little of how or why they occur. There is no mention of Tausch, Gregory, or Ganz, despite their important contributions in this area.

I can also detail several errors in this book. In the chapter on hearing, for example, Bartley suggests that combination tones appear when two frequencies are applied simultaneously but that beats appear instead when the two frequencies are close together. This is misleading. The combination tones derive from nonlinear properties of the ear, which, effectively, generate the cross products of the two applied tones; beats occur when any two frequencies are close together. They are not mutually exclusive. My major objection is that this book is a well-intentioned catalog of loosely connected facts and that the author's attempts to bring those facts together are not at all impressive. I do wish this volume were up to the high standard Bartley set for our field with his book on vision in 1941.

If a prize were given for cataloging, the winner would be Leonard Zusne. He has provided us with *Visual Perception of Form*, a monumental survey of the literature on his title topic. As the reviewer, I have been in the unenviable position of having had to read all of it at once. I do not recommend it for that purpose. However, I do believe that the serious student of form perception

should have a copy at hand. The book attempts a systematic and comprehensive review of the visual perception of two-dimensional form. It is largely successful in its comprehensiveness, but the goal of systematization was not achieved. The author claims that some degree of joining was achieved, but he confesses that the area just cannot be contained within a single theoretical framework. That is not Zusne's fault.

The historical introduction to the book is a very nice survey of ideas basic to the modern study of form perception. There are some errors. Lotze, for example, is identified as a nativist—which he was not. Zusne quotes James as having disposed of Lotze in a few words, but if Zusne would look at page 157 of the second volume of the *Principles*, he would see that James owed Lotze an enormous debt. Lotze believed that each sensation has a unique quality, which in time and with experience, comes to be identified with a place in the spatial order. It was Hering who took over the local-signs concept for the nativists.

Some other statements by Zusne are a little shocking. He asserts that "Wertheimer was the originator and 'brain' of the Gestalt school" (p. 10). This is just plain unfair. Wertheimer may have been the revered senior member of the group, but neither Köhler nor Koffka were exactly dullards. Köhler is described as a 'front man.' I believe that is being a little cavalier with the reputation of a truly great man. There are other instances of this sort of flippancy, but they are not too serious. What is far more important is that Zusne tends to emphasize the mere discoveries and specific theories of the Gestalt movement rather than its most important conceptual contribution—the denial of the constancy hypothesis.

It used to be assumed that sensations sum up with images to give perceptions. This was an assumption of linearity, because it is only in a linear system that the superposition principle applies. If the sensations were not linearly superimposable, then it would be impossible to separate them from each other during introspective analysis. Gestalt psychology denied that sensations are superimposable. Rather, it held that when diverse sensory inputs are applied simultaneously, then they inevitably transform each other. This transformation is the Gestalt. It has an existence of its own, and if it were analyzed into its constituent elements that identity would be lost. It is this principle of interaction that was the major conceptual contribution of Gestalt psychology. Today we would say with Cornsweet that the system is nonlinear.

I suppose that it is Zusne's sense of proportion with which I must disagree. He considers the various attempts to apply information theory to psychology a major breakthrough. Important, yes. Breakthrough, no. Every engineer who read Shannon was willing to apply information theory to psychological problems as soon as he had heard of them. No major contribution seems to have really come of it as yet.

The volume's short section on neuroanatomy and neurophysiology could have been left out. Some of the most elementary facts are wrong. The old superstition that there is a one-to-one connection between cones and bipolars is repeated, when in fact there is considerable divergence from the cones to the bipolars, and other complexities as well. The notion that there is a point-to-point map of the retina on the occipital cortex is also repeated here, despite the acknowledgement of the existence of units with complex

receptive fields. Somehow the inherent conflict between these ideas is not evident to Zusne; although Dodwell, in his book, recognizes it clearly, as we shall see.

Zusne also talks about the uncorrelated tremor in the two eyes and asserts that this undermines the idea of corresponding points. We know from Ogle's work on the empirical horopter that the tremor is well within the error in localizing corresponding points. One surprising omission in this otherwise encyclopedic work is that of many of the very important ideas about the mechanisms underlying the formation of Mach bands. This is clearly relevant to discussions of contour inhibition. Neither Bartley nor Zusne manages to do as well with this subject as does Cornsweet. Also, Zusne does not mention the important observation that figural aftereffects occur with stabilized images in his account of the Osgood-Heyer theory of figural aftereffects.

Although Zusne is more inclined to express his own ideas and criticisms than is Bartley, he is not always very convincing. In disputing the Pitts and McCulloch model, he asserts, for example, that seeing goes on even when the alpha rhythm cannot be detected in the EEG. This is not a valid criticism, however, because there is a great deal of activity at about 10 Hz, the scanning frequency postulated by the model, even when so-called beta rhythms predominate in the complex wave of the EEG. He also asserts that fixations are concentrated along contours and do not tend to dwell at the centers of gravity of forms. This may be true for large forms or on initial fixation, but Richards and I have reported extensive evidence for central fixation of forms that can be viewed in a glimpse. Hochberg has also reported central fixations of forms. Contours are scanned in very large forms where sequential viewing is required to get some idea of what they look like. This is not to say that I support the Pitts and McCulloch model, admirable though it may be. It is vulnerable on other grounds, however.

Zusne provides us with a fairly extensive discussion of the pattern-recognition models proposed by Deutsch, Sutherland, and Dodwell. There will be more to say about Dodwell's model when his book is reviewed below. For the moment, I would only like to mention that the juxtaposition of these models is quite useful. I do object, however, to Zusne's statement that the Deutsch and Dodwell models are "only half-baked" (p. 90). This is not only impolite but a bit unfair, since all models in this field are incomplete. It is not a crime to suggest incomplete theories. I like Zusne's observation that movement, size, and texture play an important role in the species-characteristic responses noted by the ethologists but that the typical form parameters do not seem to be involved.

There are a number of unfounded claims in this book. For one, Zusne implies at the beginning of the book that efforts to quantify visual form have been successful. I beg to disagree—on the grounds of everything else in the book. Section after section describes some other proposal for quantification. Why should this be so if the problem is solved?

Another claim is that Hebb showed that we need to *learn* to perceive forms. This is simply not true. He started with the assumption that we need to learn to perceive forms and then devised a theory to show how this might occur. There is no conclusive proof that forms cannot be seen without learning. How can a form be separated from its ground in some primitive fashion and

not have an original coherence and integrity? The work of Hubel and Weisel can be interpreted to mean that the analysis of form features may well be inborn, whereas deprivation serves to destroy this innate capability. This notion of learning to perceive demands a clearer and less ambiguous formulation than it has yet received. Learning occurs, but the dimensions along which it makes itself evident are not apparent. When we first see a Street figure, it looks like an ambiguous set of blobs. Once it is 'recognized,' it looks different in some way. But if you ask the observer to specify how it looks different, he will be unable to do so. In some way, the form is unchanged by learning. We have yet to define *how* it is altered through experience before we can postulate models to explain it. Jastrow's duck-rabbit figure is a good illustration of this profound problem.

There are many other sections dealing with fidelity of vision, developmental aspects, ways of measuring form, and form discrimination in animals, among others. All of this is extremely valuable and interesting. I do feel that Zusne's book is a useful guide to this extensive literature, but the reader had best rely on his own reading of the primary sources for a fully accurate evaluation. Zusne did the job of getting us close to the 2,600 cited references, but it is just asking too much of one man to expect a completely satisfactory evaluation of so disjointed a field.

One thing which is missing in Zusne's book is the spatial-frequency analysis of visual form. This use of Fourier techniques occurs in some automatic pattern-recognition systems. The technique is a conceptually useful alternative to the sequential analysis of discrete events so common in models of form perception. It serves, at least, to alert us to the fact that our theories may well be dictated in part by our methodological storehouse. I know only too well that if a scientist is knowledgeable about statistics, he is likely to ascribe, say, the variations over time in some wave phenomenon to the variance in the wave. If he knows Fourier methods, he will say instead that the variations are due to nonlinear interactions of a basic waveform with extraneous noise, so that the wave becomes modulated. Both views are correct, but one may be more suggestive in terms of underlying mechanisms than is the other. So we may be captivated in this field by methods popular in other disciplines. As we shall see, Dodwell may also be a victim of his own prior exposure to digital logic.

One assumption made in this field is that form recognition occurs regardless of the position of the form on the retina. For some reason, Zusne missed the paper by Wallach and Austin which showed that in some ways the recognition of a form depends on its retinal location. I believe that we had best get facts like this straight lest we try to develop theories capable of doing too much.

Peter C. Dodwell's monograph *Visual Pattern Recognition* is a very personal work. As such, it is much more interesting than Zusne's massive project. Whereas Zusne is for reference, Dodwell should be read in two or three large gulps.

This book has 11 chapters. The first chapter is a general introduction to models of form perception and their role in psychological theory. There are two things in this chapter that I want to comment on. First, Dodwell makes

the same error that many others have made in their discussions of Gestalt psychology. He describes the concept of isomorphism as a one-to-one mapping of retinal events upon the visual cortex. This view is not unique to the Gestalt psychologists but is accepted even today, as we have seen, by Zusne and many others. I think that Dodwell has forgotten that one of Köhler's most important ideas was that of *psychophysical* isomorphism. The word 'psychophysical' is very important. Köhler meant that there is a topological correspondence between experience and what occurs in the brain. This extremely important, though often-overlooked concept was revolutionary in its day. By contrast, the other meaning of isomorphism is trivial. Gestalt isomorphism is a 'ticket' that allows the investigator to travel back and forth between the brain and reports of experience. If a one-to-one correspondence exists, then one must be able to say something meaningful about brain events if he gets an accurate phenomenological account of perceptual events. This is what allowed the development of the field theory of figural aftereffects. It is amusing to note that Dodwell himself, later in his book, actually accepts the concept of psychophysical isomorphism. On page 118, for example, he asserts that "so far as models for the visual system are concerned, one need only postulate a one-to-one correspondence between an event in the model . . . and overt responses. . . ." However, it is just as plausible to assert that there is a many-to-one relationship between brain events (and therefore events in models) and experience. This would seem to rule out *prima facie* conclusions about brains based on a simple analysis of response or experience.

The second thing of interest to me in this, Dodwell's first chapter, is a small section on modular nets capable of carrying out logical operations such as those encountered in the propositional calculus. This, as we all know, but as Dodwell fails to make clear, is identical to the binary arithmetic that underlies digital computation. Perhaps this very powerful tool may prove too seductive to resist. But a more circumspect attitude may be in order, since there is no guarantee that God is a logician of the human sort. He might have constructed brains to operate in accord with other principles of computation. Dodwell, however, believes that logic nets are basic to an understanding of psychological processes and (in my opinion) does some very interesting, though probably incorrect, things with them.

His second chapter introduces several models of pattern recognition. This chapter is a must for people interested in form perception. In addition to Deutsch's and Sutherland's very clever approaches to this problem, Dodwell describes his own related model in some detail.

Dodwell's computing network outlines how spatial information may be mapped into temporal patterns of pulses, which patterns, he asserts, are uniquely related to shapes, regardless of their orientations or positions on a receptive surface. There are several objections one might raise. Apart from Dodwell's own disclaimer, there is no evidence this network is not highly vulnerable to noise. Secondly, his model does not allow the preservation of positional information. I know, for example, when a form is above or below my own line of sight. This as well as perceived orientation is lost in the transformation process. Also, the time delays occurring at the nodes in the network that receive contour information may well build up so rapidly that it could take a very long time to generate the temporal code corresponding to

a complex pattern. Reasonable biological time constants, such as those encountered at synapses, can range from 10 msec to upward of 100 msec. Many such delays generated by complex shapes would result in a total system time constant on the order of seconds. Yet form perception is more rapid than that.

A final objection is that the time of presentation of a stimulus relative to other stimuli may not be perceived if the temporal code takes substantial time to be generated. Yet temporal order, which is not recognized in Dodwell's model, is sometimes perceptible with an accuracy of somewhat less than 100 msec. It is therefore difficult to evaluate the model unless some actual numbers are applied to it. But this is an advantage of such modeling: one can apply numbers and then perform experiments to see if what may be a reasonable logical structure can actually perform the job assigned to it. I might also mention that unit 'jj' should not be connected the way it is in the schematic of the recognizer portion of the model shown on page 29.

Despite these criticisms, Dodwell need not feel defensive about his work—as he seems to when he compares it with others. It simply needs more work to find out how good it is. It works in part with rats. We want to know how much further it may be pushed.

Dodwell's third chapter is a nice review of recent neurophysiological work. I recommend it to the student who wants to have a simply written account of all of the exciting developments in contour representation in the nervous system and related phenomena. Similarly, his next two chapters cover behavioral evidence on contour coding, such as the important work by Sutherland and computer simulation respectively. It is unfortunate that Dodwell does not describe attempts at simulation of his own model.

For me, the book warms up in the seventh chapter, which deals with binocular vision. Dodwell is devoted to fusional theories of stereopsis. I do not believe that his treatment of alternative points of view is very fair. He distinguishes between suppression theories and fusion theories. Moreover, he aligns suppression theorists with the 'bad guys,' the nativists; whereas the fusion theorists are empiricists. It seems that Dodwell's evidence for fusion theory is that disparate images appear to have their visual directions shifted toward each other. This is Werner's displacement and von Tschermak-Syssenegg's allelotropia.

Displacement is probably due to unnoticed vergence changes. This is the conclusion that Pitblado reached in his dissertation and of Ogle in explaining the Hofer effect. The demonstration of displacement with Linschoten's figure can also be shown to be due to fixation at a point other than that which one is trying to maintain while looking at the stereogram. A few nonius markers over the contours of Figure 7.8 on page 133 show clearly that when one tries to fixate as requested in the text, he actually assumes another position. A really good observer can see what Dodwell says he cannot see—diplopic images.

My own reaction here is that fusion theorists have been driven to this one single demonstration of sensory fusion because most other efforts have been shown to be explicable in other terms. But even if this single demonstration worked out, fusional theories per se cannot explain the fact that there are countless instances of stereopsis where phenomenal fusion does *not* occur. Moreover, partial fusion or displacement is not necessary for the occurrence of

stereopsis. At best, one can claim with Helmholtz that fusion does not occur phenomenally but that a similar computation may (and Helmholtz really meant *may*) occur at a level that never directly reaches awareness. Helmholtz was not a fusion theorist. I do not believe that he was a nativist either.

Suppression theories break down into two classes. First, some theorists believe that suppression accounts for the singleness of vision but that it does not account for the depth computation itself. Others suggest that alternation of suppression may be involved in generating depth. Neither view is necessarily associated with a position on the nature-nature question.

Dodwell's chapter on adaptation of the visual system describes a set of mathematical transformations of forms, transformations which are like the deformations we see through prisms. Dodwell believes that adaptation occurs only when such conformal transformations are approximate descriptions of the deformation introduced by the rearrangement procedure. Otherwise, he largely accepts Held's notions about the basis of perceptual adaptation. He relies heavily on the apparent shifts in visual direction associated with fusion to support his claim that adaptation is visual rather than proprioceptive. I do not believe that fusion is a good substitute for direct evidence that Harris is wrong.

That evidence is still not available. Moreover, the fact that patients with squint do not adapt to their condition, even when the strabismus is relatively mild, cannot be explained away. Clearly, more evidence is needed. It might be noted that this topic has lost some of the vitality it had just a few years ago. It might be that the large number of seemingly conflicting reports and the inherent complexity of adaptation processes tend to discourage experimenters, who seem to have used up all of the simple approaches to the problem. Perhaps Dodwell's approach will serve to stimulate some more research.

Despite all of these negative statements, I genuinely like Dodwell's book. It is obvious that its author is a creative and intelligent scientist. The book is suggestive of many good ideas and should be read by all specialists in perception. The one thing that must be done, as work in the areas of pattern recognition continues, is to define more carefully what it is that a recognizer learns. This is too ambiguously stated by both Dodwell and Zusne. Also, it must be noted that both of these writers seem to believe that the organism must get some kind of a picture of an object into its head—and that the organism must then have a means for weighing the picture to determine if that has been seen before. But do we have to have a 'picture in the head' to perceive? Perhaps the picture is an illusion. Perhaps we have this illusion because when we report, when we are introspective, we are actually constructing a pattern that develops over time from small segments of processed input. It is this set of sequentially developed responses occurring after the fact of presentation that may represent the pattern we are supposed to have recognized. I believe that we must take seriously the idea that the problem of perception does not end with the projection of an image on some inner screen. We need to know more about the serial processes underlying the complex operation of perceiving, which takes place over time.

I have left John C. Baird's book on the *Psychophysical Analysis of Visual Space* to the end. This volume is also mislabeled. It does not deal with the general problem of visual space but largely with the topic of judged frontal

size. Moreover, it does this over a rather restricted domain of conditions. It leaves out all of the kinetic factors that influence space and size perception. Nor is it specific about the visual cues, which are simply subsumed under the topic of stimulus uncertainty. Thus, Baird holds that when a rich environment is present, then there is a great deal of information present to influence judgments. He distinguishes between high-, moderate-, and low-uncertainty environments. There is no real theory in the book, since Baird clearly eschews any effort to deal with the mechanisms that might explain the phenomena he describes in rich detail. Nor does he want to bother with the really great problems of local signs, depth perception, and the metric of visual space.

His interests lie elsewhere. When an object is placed at one distance and a comparison object is placed at a variable distance, for example, it is possible to get a subject to adjust the size of the comparison object at each of its distances until it appears to match the standard object. The ratios of the angles subtended by the two objects when a match is achieved can be plotted onto log-log paper against the ratios of the sizes of the objects that would give actual null matches to give a straight-line function. The slopes of these power functions vary with instructions, degree of stimulus uncertainty, and other factors. Baird describes such curves in copious detail. And since clusters of these functions resemble the wings of a butterfly, he derives something he labels the butterfly model. But we do not have a model here in the sense that Dodwell uses the term. There is no real effort to elucidate how or why slope variations occur. It is all very empirical. I really do not know why the term 'model' is used at all.

Baird does not try to deal with some of the classical problems in this field. Why, for example, do distance judgments not correlate with convergence angle, whereas size judgments do vary, as though the object were moving in distance appropriately with the vergence change? Nor does he consider the issue of the ways in which cues interact with each other, particularly the way in which Gibson has rephrased this entire problem. What we have is another catalog where the facts are right but the only unifying themes are such things as good fits to the magical power function. It seems to me that one of the major problems of psychology is to show why power functions are such good summaries of psychophysical data.

If the patient reader would like to have a reasonably complete account and classification of many of the facts in this highly specialized field of static frontal-size perception, he could turn to this monograph for a reliable treatment. If he is looking for profound insight and overt enthusiasm to maintain his interest, he may have to look elsewhere. The book is much more difficult than it need have been. I think that Baird would do well to read Cornsweet before he undertakes a revision of this work.

Lloyd Kaufman, *New York University*

Fundamentals of Learning and Motivation

By Frank A. Logan. Dubuque, Iowa: Wm. C. Brown, 1969. Pp. xviii, 226.

Fundamentals of Learning and Motivation was designed to introduce the beginning student to the basic principles of learning and motivation, and to

illustrate the application of these principles to the understanding of "personally important responses" (p. 2). The principles Logan describes have been derived largely from the study of lower animals, and the book reaffirms the once dominant but now widely questioned view that we can most efficiently arrive at an understanding of a wide variety of interesting human behaviors by first understanding the behavior of lower animals. For those who reject this assumption, the book is at best a partial introduction to learning and motivation.

Logan's attempt at a "coherent oversimplification" (p. xvi) is appropriate to an introductory text and has resulted in a very readable book. The treatment of classical conditioning, instrumental training, and motivation is in the tradition of Hull and Spence, with the chapter on "A Scientific Theory of Classical Conditioning" the clearest example of this bias. Logan is, however, very open about his biases, and the book can be used effectively by those who do not share them.

This reviewer would like to see expanded coverage given to several areas. First, the discussion of classical conditioning would be improved by the addition of Rescorla's description of procedures in terms of the relative probabilities of reinforcement in the presence and absence of various conditioned stimuli (R. A. Rescorla, *Psychological Review*, 1967, 74:71-80). This description, which has led to an extension of the scope of research in classical conditioning, sets the stage for discussion of excitatory and inhibitory conditioning and provides a coherent framework within which the proliferation of new paradigms can be placed.

Second, the student's attention should be called to the fact that the outcomes of many experiments in conditioning and learning depend on the particular stimulus-response-reinforcer combinations selected, as, for example, in the acquisition of learned food aversions, species-specific defense reactions, autoshaping, and so on (R. Bolles, *Psychological Review*, 1970, 77:32-48; P. Brown and H. Jenkins, *Journal of the Experimental Analysis of Behavior*, 1968, 11:1-8; J. Garcia and F. Ervin, *Communications in Behavioral Biology*, 1968, 1:389-415; M. E. P. Seligman, *Psychological Review*, 1970, 77:406-418). These data are important because they suggest limits on the breadth of our inductive leaps, for example, across species and responses, thereby suggesting that a greater variety of models than has heretofore been proposed will be needed to account for the diverse phenomena of conditioning and learning.

Vincent M. LoLordo, *University of North Carolina*

N-Person Game Theory: Concepts and Applications

By Anatole Rapoport. Ann Arbor: University of Michigan Press, 1970. Pp. xi, 331. \$3.95.

What can one say about an area of inquiry that has almost completely changed from its original domain of interest? Game theory started as the study of how to make effective decisions under conditions of conflict. Now N-person game theory is a branch of mathematics, and as Rapoport shows, quite a lot can be said about the product and process of this change.

Rapoport for many years has been a thoughtful and consistent writer about

models and human behavior. His most recent contribution is a less readable and introductory sequel to its companion *Two-Person Game Theory*. The area is difficult to communicate to behavioral scientists because it is now so mathematical; its contributors have been applied mathematicians. How many know of Shapley, Harsanyi, Nash, Shubik, or Von Neumann? Rapoport shows that it is important to know this work, if not the people, because it elucidates the logic of conflict. The logic follows at three levels of abstraction.

At the level closest to the phenomenon of the conflict of interests is *extensive form*, a specification of decisions by whom and when. These decisions can be represented by a 'tree' whose nodes are particular decision opportunities and whose branches are actual decision alternatives. At the tip of each path through the tree is an outcome to each decision maker. If one can construct the tree—a difficult empirical problem, ludicrous even in chess—one can try to find the best path through the tree. However, the best path is more easily found by representing each path as a single entity called a strategy and then focusing on the outcomes.

The second level, *normal form*, tries to sidestep specification of all the decisions and focuses on sets of strategies and outcomes. For two persons, the organizing tool is a matrix. The rows of the matrix are the strategies of one party, the columns the strategies of the second party, and the cells are the outcomes. For three persons, one would need a cube and for N persons an N -dimensional figure, a most cumbersome communication device. Therefore, Rapoport uses the two-person matrix to summarize the analyses at the level of normal form. Essentially this becomes a review of two-person game theory, emphasizing differences between pure conflict and partial conflict situations. (Experimentalists are familiar with the matrix format, which has been used peculiarly to model empirically the formal model.)

The third level overcomes the complexity of the outcome structure for more than two persons. The organizing tool is the *characteristic function*. It is a rule assigning a single number to each of the subsets of all N persons (including the null subset and the total set, the latter being an example of the fact that half of the subsets are complements to the other half). It does this by making two assumptions. The first is that the value of an outcome to each party can be combined, or pooled, and resubdivided if necessary. The second assumption is that everyone else will get together and do the best for themselves and that therefore one's strategy will be to minimize one's loss or maximize one's minimum gain. The rule's main property is that the value assigned to a coalition is greater than the sum of the values of any subcoalition.

Rapoport covered this material in the first two chapters for two reasons. First, it was preliminary to the theoretical findings reported in the first half of the book. And second, it demonstrated how the original question of effective decisions had become lost. In fact, strategies and decisions had disappeared and been replaced with challenging analyses of a rule that specified the value of a coalition.

The theoretical results show that, under various assumptions, a coalition will bring the group the best outcomes. Also, some of the results discuss rules for distributing the outcomes among the coalition members. In order to do this, the mathematics of set theory is used. For those who need to brush up,

the introductory chapter gives a brief course. With these tools, one can learn what is an imputation, core, kernel, or bargaining set.

The last half of the book applies the theoretical results to analyses of markets and legislatures, both of which are reasonably defined by the cost of a 'good' or the yea or nay of a vote. These applications, while interesting, are probably not of great importance to psychologists unless one has a well-defined voting situation in a committee.

There are several situations Rapoport does not describe that might be of interest. To the extent that one can map out a set of attitudes by an individual, or one attitude by a set of individuals, one might apply a game-theory analysis to look for various kinds of coalitions. Results might bring fresh insight and provide an alternative to the balance theorem of just two sets of attitudes (coalitions), homogeneous within a set and antagonistic between sets.

Another application might be to that of the organization of influence or communication within small or large groups. This would be an extension of the two-person game theory presently in progress on the organization of outcomes in a two-person group. In fact, if one likes to think by analogy, *N*-person game theory is a worked-out system that may provide some insights into the logic of the phenomena of your interest, and this book will help you along.

Gerrit Wolf, *Yale University*

Books received

(The books listed here have not as yet been noted in our pages. Listing here does not preclude their later review, however.)

- Abroms, Gene M., and Greenfield, Norman S., eds. *The New Hospital Psychiatry*. New York: Academic Press, 1971. Pp. x, 298. \$10.00.
- Adler, Manfred. *A Parent's Manual: Answers to Questions on Child Development and Child Rearing*. Springfield, Ill.: Charles C Thomas, 1971. Pp. vii, 111. \$5.50.
- Anderson, Michael, ed. *Sociology of the Family*. Baltimore: Penguin Books, 1971. Pp. 352. Paperback, \$2.95.
- Andrews, Lewis M., and Karlins, Marvin. *Requiem for Democracy? An Inquiry into the Limits of Behavioral Control*. New York: Holt, Rinehart and Winston, 1971. Pp. ix, 148. \$5.95.
- Arnheim, Rudolf. *Visual Thinking*. Berkeley and Los Angeles: University of California Press, 1971. Pp. xi, 345. Paperback, \$4.25.
- Arnold, William J., and Page, Monte M., eds. *Nebraska Symposium on Motivation*. Lincoln: University of Nebraska Press, 1971. Pp. xiv, 288. \$7.95; paperback, \$3.75.

continued on page 20

ANNOUNCEMENT

Research Committee on Sociolinguistics

The Research Committee on Sociolinguistics of the International Sociological Association (RCS-ISA) came to life at the time of the Sixth World Congress of the ISA in Evian, France (1966). The need had been recognized, for an organization which would assist scientists working in the area of sociolinguistics and the sociology of language, to exchange ideas and cooperate across disciplinary and national boundaries. After two years of vigorous activities and wide interest, official recognition by the Executive Committee of the ISA as an ISA Research Committee was achieved.

The Committee aims at disseminating information on sociolinguistic activities, stimulating discussion of theoretical fundamentals as well as ongoing work and organizing regional, national, and international conferences where closer contact and intensive exchange among concerned scientists can take place. The Committee's primary instrument of communication, the *Sociolinguistics Newsletter* is now published four times a year and is currently being received by approximately 600 scholars from 50 countries.

Annual dues, including subscription to the *Newsletter*, are: Regular, \$6.00; Student, \$4.00; Institutional (libraries and organizations), \$10.00. Back issues (1967, 1970) are available for \$1.00 and \$2.00 respectively.

Requests for sample copies of the *Newsletter* and applications for individual or group membership should be sent to

Evangelos A. Afendras
Secretary-Treasurer, RCS/ISA
Social Science Research Institute
1914 University Avenue, #101
University of Hawaii
Honolulu, Hawaii 96822
U.S.A.

Now, in one three-volume encyclopedia...

The first comprehensive standard reference work dealing with all facets of psychology today

Here is the most accurate, readable compendium of current psychological thought ever assembled, a practical tool for professionals and laymen alike—the new, totally up-to-date **ENCYCLOPEDIA OF PSYCHOLOGY**.

For every man and woman concerned with this fascinating and vital discipline, many advances in the field of international psychology have been a source of frustration as well as excitement. New experiments and insights have undeniably changed the course of man's journey toward self-understanding. But for most readers, the results of current research and thought are inaccessible—available only in specialized publications, and often in a foreign language.

Now, in response to the need for a single comprehensive reference work in the field of psychology, Professor H. J. Eysenck has collaborated with hundreds of international specialists to present the three-volume **ENCYCLOPEDIA OF PSYCHOLOGY**—a major tool for the understanding of current psychological knowledge.

COMPLETE, UP-TO-DATE, READABLE

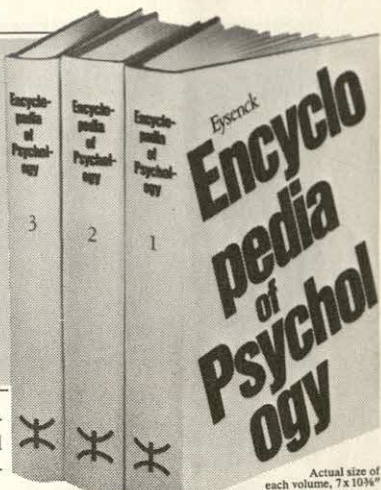
The new **ENCYCLOPEDIA OF PSYCHOLOGY** contains over 1200 pages and more than 120,000 words, with 5000 entries and 300 major articles by over 300 distinguished contributors. Each major entry offers a full definition of terms, an historical overview, a rundown of methodology, discussion of research and scientific controversies, and descriptions of the varying approaches of different psychological schools and related disciplines, as well as a complete bibliography. The **ENCYCLOPEDIA OF PSYCHOLOGY** provides both essential background information and a summary of leading international opinion on current issues. In

addition, a thorough system of cross-references enables the reader to see the key concepts of psychology in a meaningful context.

INTERNATIONAL REFERENCE WORK

The **ENCYCLOPEDIA OF PSYCHOLOGY** is the first venture of its kind in this field. It is being published in five languages simultaneously, and internationally known authors from twenty nations have participated in its preparation. Contributors include Anne Anastasi from the United States on *Differential Psychology*; A. R. Luria from the Soviet Union on *Brain Pathology*; J. D. Teasdale from Great Britain on *Drug Dependence*; L. Shaw from South Africa on *Accident Proneness*; F. Klix from East Germany on *Cybernetics and Psychology*; and P. E. Vernon (Canada) and T. Takuma (Japan) on *Personality*. Many additional contributions have been made by scientists from countries such as Czechoslovakia, Italy, Israel, and Japan, whose studies until now have been all but completely inaccessible to American readers. In Professor Eysenck's words, "It is our hope that this Encyclopedia will play a part in making our science more truly international than it has been, by drawing attention to work done in other countries, and making people more aware of the international character of modern psychology."

By subscribing now to the **ENCYCLOPEDIA OF PSYCHOLOGY**, you may assure yourself of receiving special pre-publication prices with substantial savings. Send your order today on the convenient coupon.



300 AUTHORITIES FROM 20 COUNTRIES:

UNITED STATES: L. Berkowitz on *Aggression*, A. Bandura on *Socialization*, J. E. Birren on *Aging and Gerontology*, R. B. Cattell on *Factor Analysis*, E. R. Hilgard on *Behaviorism*, ☐ **GREAT BRITAIN:** R. Lynn on *Arousal*, I. Irene Martin on *Emotion and Stress*, A. Maxwell on *Correlational Techniques*, G. R. Wilson on *Objective Tests*, I. Oswald on *Sleep*, G. S. Spinks on *Psychology of Religion*, S. Rachman on *Aversion Therapy*, ☐ **SOVIET UNION:** A. R. Luria on *Aphasia*, A. N. Sokolov on *Thought and Inner Speech*, ☐ **FRANCE:** J. Leplat on *Occupational Psychology*, ☐ **HUNGARY:** F. Illyes and A. E. Lanyes on *Educational Guidance*, ☐ **GERMANY:** W. Schraml on *Clinical Psychology*, L. Pongratz on *Mind-Body Problem*, H. Kirchhoff on *Expression*, ☐ **CANADA:** D. E. Berlyne on *Experimental Aesthetics*, A. B. Kristofferson on *Attention*. And many many more.

H. J. Eysenck, Editor of the **ENCYCLOPEDIA OF PSYCHOLOGY** is Professor of Psychology at London University, and Director of the Institute of Psychiatry at Maudsley Hospital. He has been hailed by *The London Times* as "one of the greatest living psychological experimenters."



HERDER AND HERDER
A Division of McGraw-Hill Book Company
232 Madison Ave., New York 10016

Please send _____ set(s) of the **ENCYCLOPEDIA OF PSYCHOLOGY** in three volumes at one of the special pre-publication subscription prices being offered for a limited time.

☐ M41073 M41074 M41075
Pre-publication subscription price: \$67.50. We will bill you \$22.50 (plus small mailing charge) for each volume you receive.

☐ M41070 M41071 M41072
Pre-publication subscription price: \$64.50 (\$21.50 per volume). We will bill you \$64.50 (plus mailing charge) for the complete three-volume set when you receive the first volume. ☐ Additional savings: There will be no postage charge if your check accompanies your order.

Name _____

Address _____

Zip _____

Payment \$ _____ Enclosed _____ Please Bill

Under each purchase plan I will receive Volume I now, Volume II in May, and Volume III in September.

ANNOUNCEMENT

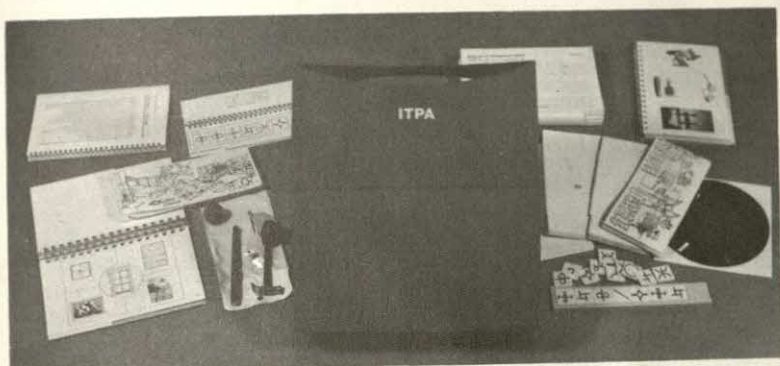
Senior Fulbright-Hays Program for 1973-1974

The Committee on International Exchange of Persons announces that applications for senior Fulbright-Hays awards for lecturing and research during 1973-1974 in over 75 foreign countries will be accepted in the spring of 1972. Specialists in the field of psychology who are U.S. citizens and have a doctorate or college teaching experience are invited to indicate their interest in these 1973-1974 awards by immediately completing a simple registration form, available on request from

Senior Fulbright-Hays Program
2101 Constitution Avenue
Washington, D.C. 20418

Registrants will receive the detailed announcement of available awards as soon as it is issued in the spring, in time to consider the possibilities and to apply before the closing date. July 1, 1972, is the deadline for applying for research awards, and it is the suggested date for filing for lectureships.

Senior Fulbright-Hays awards generally consist of a maintenance allowance in local currency to cover normal living costs of the grantee and family while in residence abroad, and round-trip travel for the grantee (transportation is *not* provided for dependents). For lecturers going to most non-European countries, the award includes a dollar supplement, subject to the availability of funds, or carries a stipend in dollars and foreign currency, the amount depending on the assignment, the lecturer's qualifications, salary, and other factors.



ILLINOIS TEST OF PSYCHOLINGUISTIC ABILITIES

Revised Edition

Samuel A. Kirk, James J. McCarthy, and Winifred D. Kirk

The Revised ITPA is an individually administered diagnostic test which taps significant abilities and disabilities in psycholinguistic functions in children between the ages of two and ten. It evaluates abilities in three dimensions:

- ☐ channels of communication (auditory-vocal and visual-motor)
- ☐ psycholinguistic processes (receptive, organizing, and expressive processes)
- ☐ levels of organization (automatic and representational)

Norms are presented for each of 12 subtests and in scaled scores. In addition, the test as a whole can yield: (1) a Composite Psycholinguistic Age and a Psycholinguistic Quotient, and (2) an estimated Binet mental age and an IQ. The complete kit includes directions for administration and scoring, 2 picture books of test materials, newly designed chips, tray, and pictured sequences for the Visual Sequential Memory Test, dispensable picture strips and scoring tissues for the new Visual Closure Test, 6 objects for the Verbal and Manual Expression Tests, a 33 $\frac{1}{3}$ -rpm instruction record, and 25 record forms. \$43.50.

A FILMED DEMONSTRATION OF THE ITPA

Produced by Winifred D. Kirk and Samuel A. Kirk

Sixteen mm. black-and-white sound film (43 minutes). \$360.00

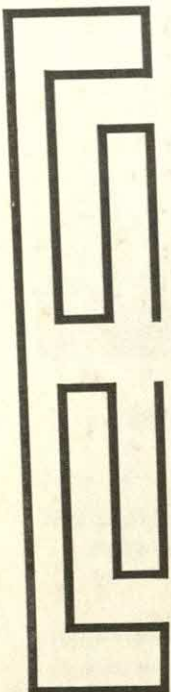
THE DEVELOPMENT AND PSYCHOMETRIC CHARACTERISTICS OF THE REVISED ITPA

John N. Paraskevopoulos and Samuel A. Kirk

Describes the model for and rationale of the tests. Paper, \$5.50.



UNIVERSITY OF ILLINOIS PRESS Urbana · Chicago · London



SEMANTICS

An Interdisciplinary Reader in
Philosophy, Linguistics and Psychology

Edited by DANNY D. STEINBERG
& LEON A. JAKOBOVITS

An inquiry into the nature of a variety of semantic problems, bringing together recent papers from three disciplines: philosophy, linguistics and psychology. Each section is introduced by a specialist who summarizes and evaluates developments. In the *Psychology Overview*, Danny D. Steinberg describes the division in psycho-linguistics between those who believe that meaning is derived from behavior and those who hold that meaning is essentially rationalistic. Both parties are represented in this section.

\$16.50

Cambridge University Press

32 East 57th Street, New York, N.Y. 10022

Handwritten notes in the top left corner:
"Pocket 2
Issue to the
Librarian 2
25/9/72"

THE AMERICAN JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys
William E. Kappauf

M. E. Bitterman
Benton J. Underwood

-
- LILA GHENT BRAINE. A Developmental Analysis of the Effect of Stimulus Orientation on Recognition 157
- ROBERT M. SCHWARTZ and MICHAEL S. HUMPHREYS. Examinations of the Category-Recall Function 189
- WALTER C. GOGEL and RICHARD D. STURM. A Test of the Relational Hypothesis of Perceived Size 201
- MICHAEL MAUS. Transfer in the A-B', A-B Paradigm as a Function of Anticipation-Interval Length 217
- C. F. STROMEYER III. Contour-Contingent Color Aftereffects: Retinal Area Specificity 227
- DAVID G. ELMES, WILLIAM I. GREENER, and WILLIAM C. WILKINSON. Free Recall of Items Presented after Massed- and Distributed-Practice Items 237

continued on back cover

Editors Lloyd G. Humphreys M. E. Bitterman
William E. Kappauf Benton J. Underwood

Book Review Editor Julian Hochberg

Consulting Editors William Epstein Robert Fox Wendell Garner
Lawrence E. Jones John Krauskopf Michael Rashotte Allan Wagner

Index Editor Rand B. Evans Necrology Editor E. R. Hilgard

The *American Journal of Psychology* was founded in the interest of general experimental psychology and publishes reports of original research, shorter notes and discussions, and pertinent book reviews. Manuscripts of articles and notes may be submitted to Drs. Humphreys or Kappauf (University of Illinois at Urbana-Champaign, Champaign, Illinois 61820), Bitterman (University of Hawaii, Honolulu, Hawaii 96822), or Underwood (Northwestern University, Evanston, Illinois 60201). Books intended for review should be sent to Dr. Hochberg (Columbia University, New York, New York 10027).

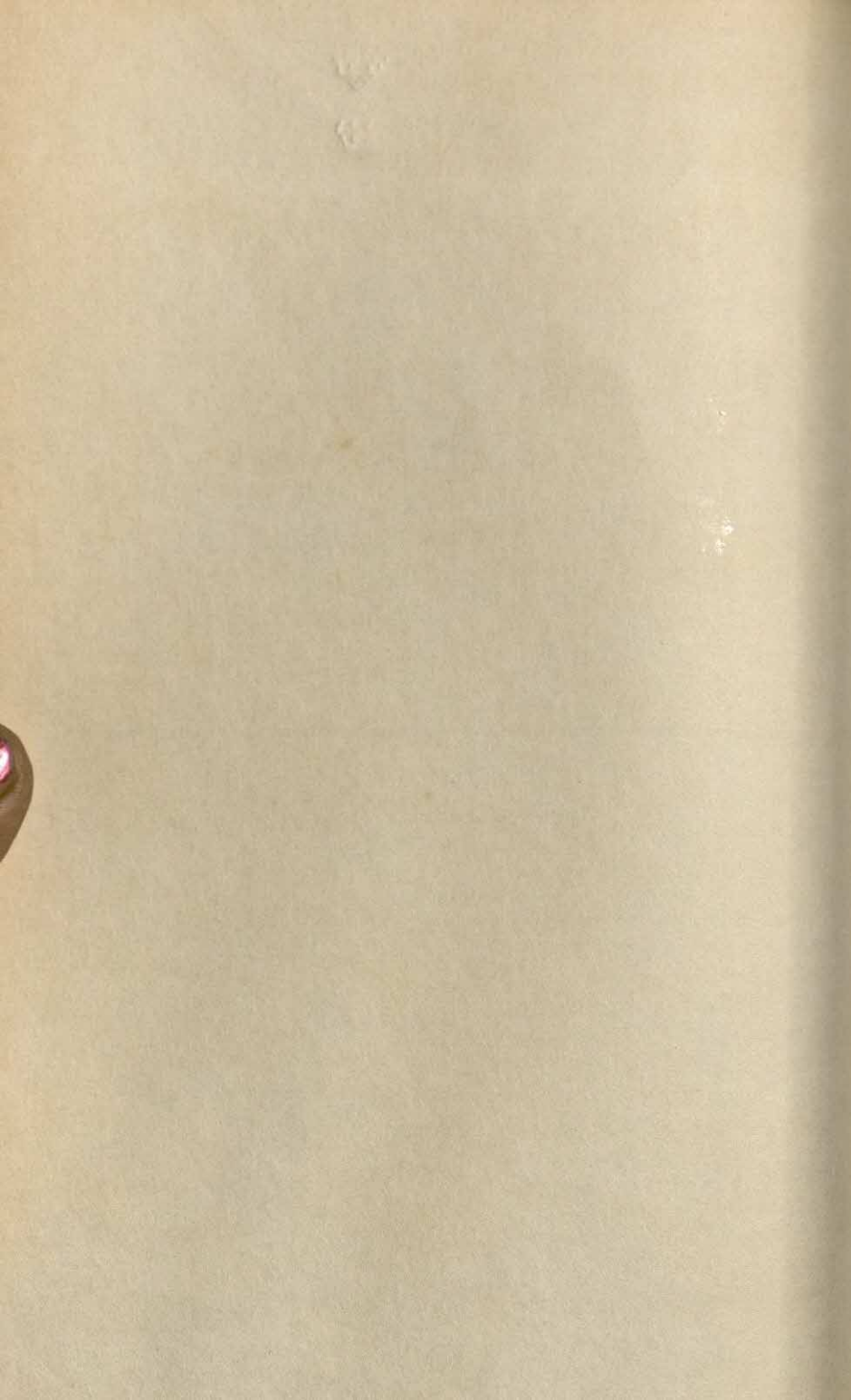
Authors of articles may order offprints at the time they return galley proof. Other contributors are sent tear-sheet copies as available unless extras are ordered when manuscript is accepted.

Contributors are asked to follow the style described in the *Publication Manual* of the American Psychological Association. Keep abstract and title quite brief, use footnotes only when essential, avoid all but the most common abbreviations (2 mm, 4 deg, 60 db, 15 sec; *but* subjects, experimenter, paired-associate learning), and use the tables and figures to supplement rather than repeat the text or each other; a well-designed table or figure can provide more information than many words of text. If you elect to omit from your manuscript any information that is critical to the evaluation of your data (analyses of variance summary tables, means, variances, correlation and factor matrices), provide two copies of that material for the editors' and their consultants' use in reviewing the paper.

The subscription price is \$10.00 a year (foreign \$10.50); single copies cost \$2.75. Back numbers and volumes may be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which all business inquiries (subscriptions, advertisements, changes of address) should be addressed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

GERALD M. MURCH and JOY HIRSCH. The McCollough Effect Created by Complementary Afterimages	241
MARI R. JONES and JAMES R. ERICKSON. A Demonstration of Complex Rule Learning in Choice Prediction	249
MORTON W. WEIR. Probability Performance: Reinforcement Procedure and Number of Alternatives	261
HERSCHEL W. LEIBOWITZ, CHARLES GRAHAM, and MICHAEL PAR-RISH. The Effect of Hypnotic Age Regression on Size Constancy	271
JAMES W. HALL. Verbal Behavior as a Function of Amount of Schooling	277
Clarence Henry Graham: 1906-1971 (Frank A. Geldard)	291
Gösta Ekman: 1920-1971 (David Magnusson)	295
Book Reviews (Phillip Shaver, Thomas Natsoulas, M. E. Bitterman, Donald Rob-bins, J. A. Michon, Dean G. Purcell, Stephen C. Woods, Olin W. Smith and Patricia C. Smith, Jon D. Swartz, Kenneth J. Gergen, Dorothea E. Johannsen; books received)	297



587
26.9.72

A DEVELOPMENTAL ANALYSIS OF THE EFFECT OF STIMULUS ORIENTATION ON RECOGNITION

Lila Ghent Braine
New York University

Subjects of various ages were tachistoscopically shown geometric figures that had a distinguishing cue at one end, and at the other end, a 'focal' feature that determines the apparent upright for young children. The orientation facilitating recognition changed with age; a developmental change in the starting position of a serial-processing mechanism is proposed.

It is an old observation, both in psychology and in everyday experience, that the orientation of a form influences its phenomenal appearance and also its identifiability. The interpretation of such observations is complicated by the fact that there appear to be developmental changes in the observer's response to orientation—a young child may be content to look at an upside-down book, but the parent cannot read it that way.

A thorough review of the work on the orientation of form appears in a recent book by Howard and Templeton on spatial orientation in man (1966, pp. 294-348). In a theoretical interpretation, these authors propose the operation of "some internalized scanning mechanism, and that a sequential signal is produced containing all the information of both the shape and the orientation of the stimulus" (p. 319), to account for the effect of orientation on recognition. Curiously, they give no reason for invoking a sequential-processing mechanism, and the very general theoretical view offered could be served equally well by parallel processing. The point of view presented in this paper also rests on the assumption that a serial-processing mechanism operates in form perception at some level in the organization of the percept, but it further assumes that this processing has some specific directional characteristics.

Now, serial processing in vision is usually assumed to be random—to start at a focal feature and proceed randomly from one focal feature to another, as suggested by Hebb (1949). The thought that serial processing has *directionality* first arose in an attempt to explain the observation

that young children consider some nonrepresentational forms upright in one orientation and upside down when turned 180 deg (Ghent, 1961). Such judgments of orientation have been elicited in Iranian preschool children (Antonovsky and Ghent, 1964), in a range of age-groups in the United States (Pick and Pick, 1966; Harris and Schaller, 1969), with a variety of figures (Pick and Pick, 1966; Koff, 1969) and different methods of presenting the stimuli (Moeller and Goodnow, 1969; Harris and Schaller, 1969).

In that 1961 Ghent study, phenomenal uprightness seemed to be characterized by the presence of a focal feature at the top of the pattern; it was possible to predict the preferences for simple figures with the focal features defined *a priori*. (The problem of defining 'focality' is discussed in the last section of the present paper.) The judgments of orientation were interpreted as deriving from processes underlying form perception—more specifically, as indicating that serial processing of the parts of a form started at a focal feature and continued in a downward direction. This interpretation implies that recognition of a form should be facilitated by presentation of the form in the upright orientation (under conditions preventing eye movements); and recognition of tachistoscopically presented forms was indeed better when the forms were presented in the phenomenally upright position than when they were presented upside down (Ghent and Bernstein, 1961).

Ghent and Bernstein (1961) also found that facilitation of recognition by the upright orientation diminished with *age*: facilitation was marked in three-year-olds, had decreased in four-year-olds, and had virtually disappeared in five-year-olds. This apparent age change in the tendency to process downward from a focal feature constituted the point of departure for the set of experiments to be reported here.

Here, the effect of orientation on recognition of geometric figures was investigated in three age-groups: first-grade children, nursery-school children, and college students. Experiments I and II form a pair: Experiment I provides data on judgments of orientation for a new set of geometric forms, and Experiment II investigates recognition of these forms presented in different orientations to children in the first grade. Experiments III and IV likewise form a pair: Experiment III describes judgments of orientation for another set of geometric forms suitable for work with nursery-school children, and Experiment IV investigates recognition of these forms presented in different orientations. Experiment V extends the investigation of the effects of stimulus orientation on recognition to college students.

EXPERIMENT I

The first question is why the effect of an upright orientation on recognition disappears around five years (Ghent and Bernstein, 1961). There are at least three possibilities. The most parsimonious is that serial processing becomes random after four years, to have no directional component; there is nothing in the data or theory on form perception to suggest that the orientation of geometric forms should influence recognition in the absence of selective experience with the form in a specific orientation. Another possibility is that the mode of serial processing remains the same in the older child but that more complex figures are required to demonstrate the effect. A third possibility is that serial processing remains directional but that it changes around five years, in such a way that the starting point is no longer the focal feature but, rather, the top of the figure as defined by the spatial framework.

In order to investigate these possibilities, new figures had to be devised to meet three requirements: (a) that the figures be more complex—have more lines in them—than those used earlier, (b) that the feature distinguishing one figure from another be different from the assumed focal feature and at the opposite end of the figure from it, and (c) that the figures elicit highly consistent judgments of orientation in the predicted direction, namely, that the upright orientation have the assumed focal feature at the top. The first and second requirements were met simply by the appropriate design of the figures; the present experiment investigates whether the third requirement was met.

Method

The subjects were 9 boys and 14 girls attending the University of Maryland nursery school. The mean age was 5:7 years, with a range from 5:2 to 6:1 years. All the children were white; most of them came from homes of professional parents.

The stimuli were three families of rectilinear figures in which each member of the family shared a common feature (assumed to be focal) at one end, had identical sides, and differed only in the way the figure was closed at the other end, this latter portion being the differentiating feature of the figure (Figure 1). The figures, drawn in black ink on a card 4 by 4 in., were $2\frac{1}{2}$ in. long and $\frac{3}{4}$ in. wide.

The experimenter and the subject sat on adjacent chairs in front of a table, and the subject was shown a pair of cards, identical in shape but differing 180 deg in orientation with the long axis vertical. The experimenter held the cards perpendicular to the child's line of sight, at arm's length in front of him, and asked him to point to the card that was upside down. Six pairs of realistic

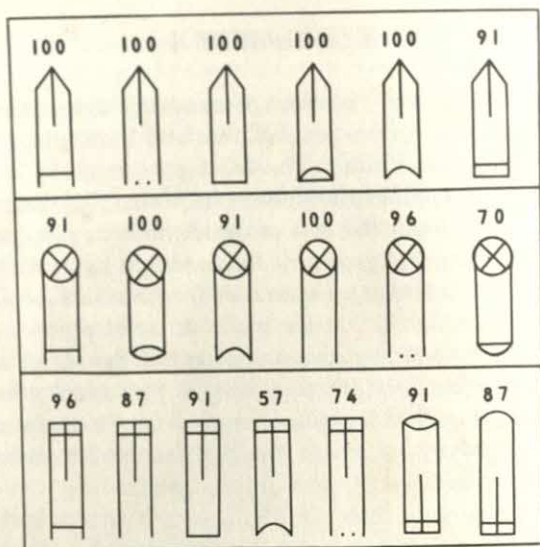


Fig. 1. Three families of complex figures in the phenomenally upright orientation; the numbers are the percentages of children judging the figures upright in the orientations shown, Experiment I

pictures were presented first, and then the 19 pairs of geometric figures shown in Figure 1. The 19 geometric forms were arranged in a fixed random order and were presented in order 1-19 to half the group, in the reverse sequence to the other half.

Results and discussion

Figure 1 shows the geometric forms in the orientations most frequently judged as upright and the percentages of children making each such judgment. The main point is that 17 of the 19 figures were considered upright when the assumed focal feature was at the top of the form. The judgment was statistically significant at the .01 level in all but three instances—the last item in the second family and the fourth and fifth items in the third family of figures. The 14 figures that more than 85% of the children judged to be upright with the assumed focal feature at the top clearly met the three requirements for further investigation of directional processing in form perception. The two instances in which the observed judgments for orientation were opposite to those predicted (the last two items of the third family shown in Figure 1) will be discussed in the section on 'focalness' at the end of this paper.

EXPERIMENT II

Let us now return to the question of the three possible bases for the change with age in the effect of orientation on recognition of shape. The forms shown in Figure 1 were designed to provide control over the part that had to be perceived in order to select the correct form from a multiple-choice array of all the members of the appropriate family. The orientation of the form that facilitated perception of the distinguishing feature would, theoretically, vary with the mode of processing used by the subject. If serial processing is random in the older child, there would be no reason to expect the orientation of unfamiliar geometric figures to have *any* effect on recognition. Alternatively, if the older child continues to process complex forms by starting with the focal feature and continuing downward, then recognition (of the complex forms shown in Figure 1) would be best with the focal feature at the top. That is, processing downward from the focal part would facilitate perception of the distinguishing feature for a figure in the upright orientation, whereas for the upside-down orientation processing downward from the focal part would lead away from the remaining lines of the figure. Finally, if processing in the older child begins at the top of the figure rather than at the focal feature, recognition would be best when the figure is oriented with the distinguishing feature at the top, since the distinguishing feature would be processed first (before the trace faded).

Method

The subjects were 21 girls and 11 boys from the junior-primary and first grades in one of the public schools in Washington, D.C., that served a population approximately half black and half white at the time of testing. The school served lower- and middle-class families. The mean age of the group was 6:9 years, with a range of 6 to 8 years. Twenty-eight children were right-handed and four were left-handed, as determined by the hand used for writing.

Ten of the figures shown in Figure 1 were selected as stimuli: the first four items in the top family, the first three items in the middle family, and the first three items in the bottom family. Each figure was presented tachistoscopically in four orientations, so that the distinguishing feature appeared at the bottom (the phenomenally upright orientation), at the top, at the left, or at the right of the figure. (The theoretical considerations underlying this experiment are relevant only to the results obtained for presentation in the phenomenally upright and upside-down orientations, but the figures were presented turned to the left and to the right to explore the effects of such orientations.)

The 40 items were arranged in a semirandom order in four different se-

quences. The child indicated the item seen by pointing to the stimulus item in a multiple-choice array of six items of the appropriate family; the array was on a large sheet of cardboard and was shown to the subject by the experimenter after the presentation of each stimulus. Four sets of multiple-choice arrays were used, each set being oriented to correspond to one of the orientations of the stimulus figures. The subjects were divided into four groups, each group receiving the arrays in a different orientation. Thus the orientation of all the multiple-choice items was the same for any particular subject but differed among groups, whereas the stimulus figures appeared in each of the four orientations for all subjects regardless of group. This design provides procedural simplicity and makes it possible to evaluate any effect that the orientation of the multiple-choice items might have on the identification of the stimulus items. Within each multiple-choice group, the subjects were counterbalanced with respect to the sequence of presentations used, and as much as possible, with respect to age, sex, and exposure duration.

Each subject was tested individually in a large room in which the light conditions were kept constant by drawing the window shades and leaving one set of room lights on. The subject sat on a chair 9 ft away from the screen, and the experimenter sat a little off to the subject's side. The stimuli were presented on a rear-projection screen by a Kodak projector (fitted with an Ilex shutter) behind the screen. A $4\frac{1}{2}$ -in. square was outlined with black tape on the screen at approximately the subject's eye level, and the subject was told to look there before the presentation of each stimulus. The long axis of the stimulus was $3\frac{3}{4}$ in. and subtended a visual angle of 1.7 deg. After each presentation, the appropriate family of multiple-choice figures was handed to the child, and he was allowed as much time as he wanted to make a choice. He was never told whether he was right or wrong, but he was encouraged in a rather general way throughout and was given a lollipop at the end of each session.

A set of 16 practice items always preceded the test items. The practice figures were two families of figures designed to have characteristics similar to those of the test figures; that is, the members of each family were alike on three sides and differed on the fourth. These figures were presented in each of four orientations in order to acquaint the child with the procedure and to give him practice in selecting figures from the multiple-choice arrays that did not necessarily match the orientation of the stimulus figures. The subject was told that the figures would appear turned in different ways and that he was to try to find the correct one on the big card, without paying any attention to how it was turned. If he had difficulty in understanding the task, several trials were run with the stimulus and multiple-choice figures present simultaneously. In general, the practice trials began with exposure durations of $\frac{1}{2}$ sec, and the exposure times were decreased as the child improved. Most subjects required more than one practice session to reach a stable and efficient level of performance, after which the test figures were introduced.

The test figures were always presented in a single session; the exposure duration was constant for a given child for the test figures but varied from child to child in an attempt to achieve a level of 50% correct for each child. Fourteen were tested at $\frac{1}{10}$ sec, two at $\frac{1}{25}$ sec, three at $\frac{1}{50}$ sec, four at $\frac{1}{100}$ sec, and nine at $\frac{1}{300}$ sec. A bimodal distribution of exposure durations required to

achieve a particular level of performance has been found previously in a group of mentally retarded subjects (Ghent, 1964), and Mayzner et al. (1968) have noted that college students seemed to fall into 'slow and fast' processors in their studies of visual-information processing.

Results

The main question is whether the orientation of the geometric figures affected the ease of recognition of the figures and, in particular, whether shifts within the vertical axis affected recognition scores. Since the multiple-choice arrays used in making a response were presented in a different orientation to each of four groups, the scores for the stimulus figures are presented separately for each group in order to evaluate any interactions between the orientation of the stimulus figures and the orientation of the multiple-choice figures.

Figure 2 shows the number of figures identified correctly when the distinguishing feature appeared at the top and at the bottom of each stimulus figure, for the four groups receiving the multiple-choice arrays in different orientations. It is clear that in every group, the figures were identified more frequently when the part necessary for recognition (the distinguishing feature) appeared at the top. These data were evaluated statistically by a two-factor analysis of variance with repeated measures over the variable of orientation of stimulus figures and independent measures over the orientation of the array figures. Only the orientation of the stimulus figures yielded a significant effect ($F = 8.17$, $p < .01$ for $\frac{1}{2}s$ df). Although one might have anticipated that a congruence of orientation of stimulus and multiple-choice figures would have facilitated recognition, the interaction was not significant. The stimulus figures oriented with the distinguishing feature at the top were recognized best when the multiple-choice array was in the same orientation, but recognition of the figures with the distinguishing feature at the bottom appeared to be facilitated equally by multiple-choice figures in the 'bottom' and the 'right' orientations (Figure 2).

When the stimulus figures were presented horizontally, the orientation of the figures did, somewhat surprisingly, affect recognition. As can be seen in Figure 3, the figures were identified better when the distinguishing feature appeared at the right side of the figure. An analysis of variance with the orientation of the stimulus and multiple-choice figures as the main variables indicated that the orientation of the stimulus figures barely reached a significant level ($F = 4.19$, for $\frac{1}{2}s$ df ; $p = .05$ requires a value of $F = 4.20$). The stimulus figures oriented with the distinguishing feature at the left were identified at a remarkably similar

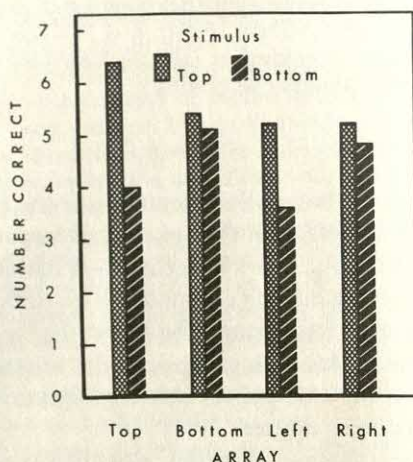


Fig. 2. The number of items identified correctly by first-graders when the distinguishing feature appeared at the top or at the bottom of each stimulus, for each array group receiving the multiple-choice items oriented differently, Experiment II

level in the different array groups, but recognition of the figures with the distinguishing feature at the right tended to be facilitated by a similar orientation of the multiple-choice figures, although, again, interaction of stimulus and array figures did not even approach a significant level.

Discussion

On the assumption that serial processing of form occurs at some level in the system, three predictive formulations were made of the nature of this processing in children above five years of age. The first formulation, that the processing occurs randomly from part to part, is not supported by the data, since recognition was better when the distinguishing feature appeared in some locations than in others. The second formulation, that processing proceeds as in young children, is clearly wrong, since even the presentation of relatively complex figures did not elicit better recognition for the phenomenally upright figures. The data support the third possibility, that serial processing starts from the top, as defined by the spatial framework, and presumably continues in a downward direction.

This last interpretation implies that a specific developmental change occurs in one of the processes underlying form perception. Before the age of five years, serial processing appears to start at the focal feature

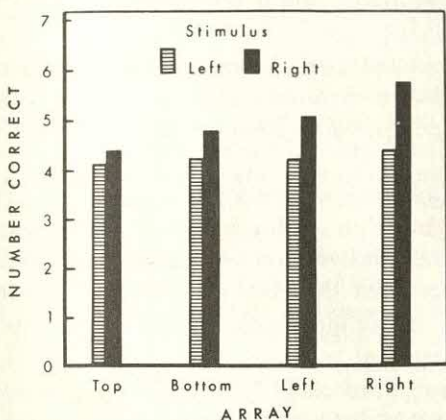


Fig. 3. The number of items identified correctly by first-graders when the distinguishing feature appeared at the left or at the right of each stimulus, for each array group receiving the multiple-choice items oriented differently, Experiment II

and to continue downward; after this age, the starting point becomes the top of the spatial framework. In more general terms, a directional-processing mechanism underlying shape recognition is influenced by a specific characteristic (the focal feature) of the stimulus in the very young child—a form of stimulus-boundedness—whereas this mechanism becomes independent of the specific stimulus input around five years of age.

The validity of this conclusion rests, however, on the somewhat shaky assumption that the results obtained with one set of figures with six-year-olds can be compared to the results previously obtained with a different set of figures with younger children (Ghent and Bernstein, 1961). The developmental change proposed here must be demonstrated by different responses at different ages to the identical set of stimuli; the next two experiments were concerned with this problem.

Discussion of the unexpected finding that horizontally oriented figures were identified better when the distinguishing feature appeared at the right than at the left will be delayed to the general discussion at the end of the paper.

EXPERIMENT III

The investigation of recognition in preschoolers using the figures used with the six-year-olds ran into problems. As might have been antici-

pated, the four-year-olds found the discrimination of one figure from another very difficult, even when the figures were presented for 1-2 sec outside of the tachistoscope. A small group of subjects was trained to make accurate discriminations under these conditions, but these same children could not perform adequately under conditions of tachistoscopic exposure.

Since the use of the complex figures with three- and four-year-olds seemed doomed to failure, a simpler set of figures was designed. It was assumed that the difficulty was due mainly to the complexity of the focal feature, and figures were designed using simple focal features—an acute angle and a convex portion—both of which were predicted to be focal features on the basis of previous work (Ghent, 1961). Two families of figures were constructed, each family with a focal feature in common and with a feature to distinguish one figure from another placed at the opposite end of each figure. These families, shown in Figure 4, differ from the complex figures used in Experiments I and II in the nature of

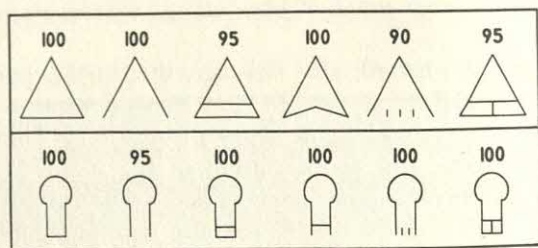


Fig. 4. Two families of simple figures in the phenomenally upright orientation; the numbers are the percentages of children judging the figures upright in the orientations shown, Experiment III

the focal feature—which is, in fact, irrelevant to the discrimination of one figure from another. The present experiment tested the prediction that the figures would be judged upright when the assumed focal feature appeared at the top; this experiment was preparatory to using the figures to investigate recognition.

Method

The subjects were 11 females and 12 males tested at a nursery school in Virginia that drew its population from a middle-class, predominantly white neighborhood. The mean age was 4:3 years, with a range from 3:3 to 5:2 years. All the subjects had spent at least one previous session with the experimenter, doing discriminations with unlimited exposure time of either the complex or the simple figures already described.

The experimenter and the child sat side by side facing a small table; the experimenter presented pairs of cards, and the child made his response by pointing to one of the cards. Members of each pair were identical except for their orientation, one member being presented in the assumed upright orientation and the other turned 180 deg (inverted). The cards were again held perpendicular to the line of the child's sight rather than placed flat on the table, since incidental observations have suggested that judgments by preschool children of 'upright' and 'inverted' are made more reliably in the frontal than the transverse plane. As each pair was presented, the child was asked to point to the one that was "upside down or looked wrong." Five pairs of realistic figures were presented first: table, flowers, house, man, and cow. Then the 12 pairs of geometric forms shown in Figure 4 were presented, the prediction being that the figures would be considered upside down when the focal feature was at the bottom. The families were alternated, resulting in a fixed order that was presented in one sequence to half the subjects and then in the reverse sequence to the other half. The realistic figures always preceded the geometric, and if a child made an error on these, he was corrected and the group of five realistic pairs was repeated before beginning the geometric figures. Eight children made one or two errors on the first presentation of the realistic figures. Five of these made no errors on the second presentation and were included as subjects, whereas the other three children made several errors on the second presentation of the realistic figures and were excluded, leaving a total of 20 subjects.

Results and discussion

The subjects considered all the figures upright with the focal feature at the top; the percentage of subjects making this judgment for each of the cards is shown in Figure 4. The high consistency of the judgments in the predicted direction indicates that this new set of figures fulfilled a main requirement for investigation of the effects of orientation on recognition.

EXPERIMENT IV

The developmental change in processing suggested earlier required that two specific predictions be met in this experiment.¹ First, recognition in three-year-olds should be better when the figures are presented in the phenomenally upright orientation—which, for these particular figures, means when the distinguishing features appear at the bottom of the figures. This prediction stems from the assumption that three-year-olds process a form by starting at the focal part and continuing downward: the presence of the focal feature at the top would be congruent with this order of processing and should facilitate perception of the distinguishing feature. The second prediction is that recognition in older children should be better when the figures are presented with the dis-

tinguishing features at the top of the cards, since processing starts at the top first in older children. On the basis of previous data (Ghent and Bernstein, 1961), it was anticipated that the change should occur around five years of age.

Method

The subjects were tested in a nursery school in Arlington, Virginia, that served a predominantly white and middle-class population. Work was begun with 42 children, but 6 children, all less than 4;4 years, did not complete the tachistoscopic procedure and their data are excluded from all computations. The youngest group had 10 children in the age range 3;1–3;11 years; the intermediate group, 10 children 4;0–4;5 years; and the oldest group, 16 children 4;6–5;1 years.

—Nontachistoscopic exposure—Because of the difficulty of the discrimination required, a training session was given first, in which the stimuli were exposed for 1–2 sec. The experimenter and the child sat side by side in front of a small table on which was placed a small stand for holding the multiple-choice cards. A series of ten easily discriminated realistic and geometric figures was presented first. A card with a single figure was placed in the hand of the child, he was allowed to hold it for 1–2 sec, then the card was taken away and a larger card containing four multiple-choice items was placed on the stand close to the child and he was asked to point to the picture he had just held. If he had any difficulty, he was allowed to compare the stimulus and multiple-choice card simultaneously early in the training session, but not later. The children were encouraged and aided to make sure that they understood the task; such help was necessary mainly for the youngest group of children.

After the set of easily discriminated forms was mastered, the set of 12 figures shown in Figure 4 was presented, each figure appearing once in the upright orientation and once in the upside-down orientation. Again, the child responded by selecting one of the four figures of the same family presented on the multiple-choice card; the figures were the same size as the stimulus figures. The four figures appeared on the multiple-choice card in a diamond configuration (at the top, bottom, left, and right corners of the diamond), which seemed to elicit fewer response biases than presenting the items in a row or a column. A different multiple-choice array was made for each stimulus card; the arrays for all subjects were presented in the phenomenally upside-down orientation for this phase of the training (primarily as a matter of convenience).

The subjects were trained to a criterion of six correct of the last eight presentations and were then considered ready to go on to the tachistoscopic procedure, which was done at a subsequent testing session one or two days later. Two subjects in the oldest group did only half the test items under conditions of long exposure, because they achieved perfect scores for both the test items and the easily discriminable forms presented first.

—Tachistoscopic exposure—A Keystone projection tachistoscope, fitted with an Ilex shutter, projected the stimuli from behind a rear-projection screen.

The subject sat directly in front of the screen, 50 in. away from it, and the experimenter sat at the child's side, in a position to watch his eyes and to place a multiple-choice array on the stand slightly to the side of the child. A small, 2-in., square was outlined with fine black tape on the screen, and the subject was told to look there before each exposure. The figure itself was $1\frac{1}{2}$ in. long, subtending a visual angle of 1.7 deg. A series of geometric figures, similar to those used during the training period, was used for practice, and they were presented in two orientations. During the practice series, an exposure duration was identified, for each child, at which approximately half the items were identified correctly, and this exposure was used for presentation of the test items. The exposure duration was constant for a given child for the test figures but varied from child to child in order to achieve a comparable level of performance in the different age-groups. In the youngest group, two were tested at $\frac{1}{8}$ sec, six at $\frac{1}{10}$ sec, and two at $\frac{1}{25}$ sec. In the intermediate group, four were tested at $\frac{1}{10}$ sec, four at $\frac{1}{25}$ sec, and two at $\frac{1}{50}$ sec. In the oldest group, one was tested at $\frac{1}{10}$ sec, eight at $\frac{1}{25}$ sec, five at $\frac{1}{50}$ sec, and two at $\frac{1}{100}$ sec. Previous studies have also found that the exposure duration required for a given task varies inversely with age in children (Ghent, 1960; Ghent and Bernstein, 1961).

The stimulus figures were divided into two groups of six figures, three from each family, each presented once upright and once inverted in a semirandom order. A group of 12 trials was presented, then testing was stopped and the child and the experimenter played together (with play dough, or crayons, or a small car) for 5 min or less, after which the child either continued with the second group of figures or left the testing situation, to return in the next day or two to complete the second group of 12 items. If a second session was required for the test figures, the practice set of items was used again. Only ten subjects, seven of them from the youngest group, required two separate sessions for the tachistoscopic presentation of the test items. The subjects in each age-group were balanced with respect to the group of items and orientation seen first. The items in the multiple-choice arrays were in the upside-down orientation for the two groups of younger subjects (an orientation which might decrease the predicted effect); in the group of older subjects, half received the arrays in the upright orientation and half in the upside-down orientation. Previous work has indicated that array orientation is not a significant variable for preschool subjects (Ghent and Bernstein, 1961).

In contrast with previous work, the subjects were rewarded for correct responses during tachistoscopic exposure of both practice and test items, in order to make the procedure more attractive to the younger subjects. A small wooden holder with four holes was placed in front of the child, and he received a marble to place in one of the holes for each correct response; a full holder of marbles was then exchanged for a small candy.

Results

The necessity for preliminary training before tachistoscopic testing provided the opportunity to compare the effects of orientation on recognition under two conditions: with a relatively long exposure duration,

and with a brief exposure duration that did not allow eye movements. According to the line of thought developed thus far, recognition should be influenced only under the condition of brief exposure, since unlimited exposure would allow the subject to move his eyes over the whole stimulus and examine all parts.

First, the results on the *long exposure*. As anticipated, the orientation of the stimulus figures did not affect accuracy of recognition at any age. The numbers of items correctly identified in the upright and inverted orientations were, respectively, 6.1 and 6.4 for the youngest group, 8.1 and 8.0 for the intermediate group, and 8.9 and 8.0 for the oldest group. (The maximum score for each orientation was 12.) It is clear that performance improved with age; this improvement may be even more marked than these scores show, because the younger children were often given more than one training session and the scores for the last session were used.

An unexpected finding was that the two families of figures were not identified equally well, the family of triangular figures being considerably easier than the family of mushroom-shaped figures for all age-groups. The numbers of items correctly identified for the triangular and mushroom-shaped families were, respectively, 6.9 and 5.5 for the youngest group, 8.9 and 7.2 for the intermediate group, and 9.9 and 8.0 for the oldest group. For the group as a whole, 26 subjects achieved higher scores for the triangular items than for the mushroom-shaped items, 1 showed equal scores, and 7 achieved higher scores for the mushroom-shaped items ($\chi^2 = 10.62, p < .01$).

Next are the results on the *tachistoscopic exposure*. Under this condition, the orientation of the stimulus figures did affect accuracy of recognition, and furthermore, the nature of the effect changed with age, as Figure 5 demonstrates. The youngest group of children recognized the upright figures (focal feature at the top) better than the same figures in the inverted orientation ($t = 3.19, df = 9, p < .02$). The oldest group of subjects, on the other hand, showed better recognition for stimulus figures in the inverted orientation (distinguishing feature at the top) than for those in the phenomenally upright orientation ($t = 5.62, df = 15, p < .001$). The children of intermediate age showed no consistent effect of orientation, presumably because the group was in the state of transition.

As in the condition of long exposure, the orientation of the multiple-choice figures did not seem to influence performance. The youngest and intermediate age-groups received the multiple-choice items only in the inverted orientation, yet recognition was not better for inverted stimuli.

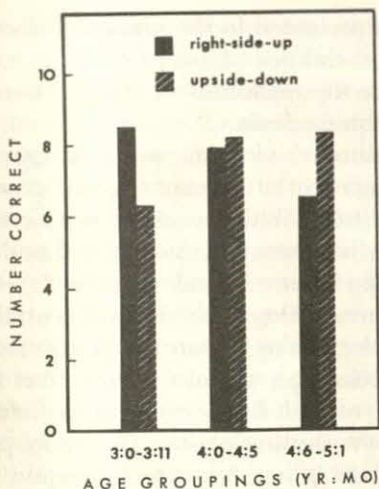


Fig. 5. The number of items identified correctly by three different age-groups when the stimuli were presented in the upright and upside-down orientations (distinguishing features at the bottom for the upright stimuli and at the top for the upside-down stimuli), Experiment IV

Results for the oldest group provide the most direct test of the effect of the multiple-choice items on recognition of the stimulus figures, since half the group received the multiple-choice items only in the upright orientation, the other half only in the inverted orientation. For the subjects receiving the upright multiple-choice items, the number of correct identifications was 6.1 for the upright stimuli and 7.8 for the inverted stimuli. For the subjects receiving the inverted multiple-choice items, the number of correct identifications was 6.9 for the upright stimuli and 8.9 for the inverted ones.

Again, a marked difference appeared in the ease with which the two families of figures were recognized. The numbers of items correctly identified for the triangular and mushroom-shaped families were, respectively, 8.8 and 6.0 for the youngest group, 9.1 and 7.0 for the intermediate group, and 8.5 and 6.4 for the oldest group. For the group as a whole, 26 subjects achieved higher scores for the triangular items than for the mushroom-shaped items, 5 showed equal scores, and 5 achieved higher scores for the mushroom-shaped items ($\chi^2 = 12.25$, $p < .01$).

Discussion

The results support the two predictions made on the basis of previous work. That is, very young children identified geometric shapes better

when the shapes were presented in the phenomenally upright orientation. And at a later age, children identified the shapes better when the relevant cue was at the top regardless of whether the figure was phenomenally upright or upside down.

As would be anticipated, this change with age appeared only under conditions in which the effect of a temporospatial processing could be expected to manifest itself. With prolonged inspection, the subjects would be able to move their eyes over the whole stimulus, and the location of the distinguishing feature should not have any effect on recognition. With brief exposures of the whole shape in central vision, the spatial location of the distinguishing feature would determine the order in which this cue was processed, and hence would affect recognition. The change with age observed with brief exposures has been interpreted as reflecting a change in the starting position of a serial-processing mechanism; processing tends to go in a downward direction, and the starting position changes from the focal feature to the top of the figure as defined by the spatial framework.

Another way of describing this change is to say that processing starts at the top and continues downward in both young and older children but that the definition of the phenomenal top changes with age. For the very young child, the top is defined by the feature characterizing the top of the figure, whereas for the older child, the top is defined by the spatial framework. Such a formulation has the attraction of simplicity, since the processing rule would remain the same at different ages but the definition of one of the variables would change with development.

The problem with this last formulation is whether there is an unambiguous way to designate the phenomenal, or apparent, top of a figure. For example, if asked to point to the top of an inverted picture of a person, one might point to the head portion (top of the figure defined by a feature) or to the feet portion (top defined by the spatial framework). The decision to make one or the other judgment would depend more on a cognitive assessment of the requirements of the task than on an appearance of 'top.' (It is assumed that these alternatives are available to the older child and adult but that the young child can judge only 'top defined by a feature.') Because of the difficulty in defining the phenomenal top of the figure for the older child, the developmental change will be described as a change in starting at the focal feature to starting at the top as defined by the framework.

An unexpected finding of the present experiment was the difference in difficulty level between the two families of figures under conditions of brief and long exposure durations. Although the distinguishing features

were almost identical in the two families and although the perimeters of the families were comparable, there were two differences between the families that could influence discriminability. One was that the proportion of the perimeter occupied by the distinguishing feature was much greater in the triangular family; the other was that only one turn, or angle, appeared in the common feature of the figures in the triangular family (the 'family feature'), whereas two angles appeared in the family feature of the mushroom-shaped figures. In an unpublished study designed to investigate whether the difference in the family feature was sufficient to produce a difference in recognition, new figures were made in which the distinguishing features, or relevant cues, were identical in size and shape for both families.² In a group of 26 three- and four-year-olds, the triangular figures were identified more often than the mushroom-shaped ones ($\chi^2 = 4.65$, $p < .05$). Again, recognition was poorer for the figures with more turns in the family feature, even though the family feature was irrelevant to identification under the conditions described. Similarly, it will be recalled that the figures used in Experiment IV represent a simplification of an irrelevant cue (compare the family features in Figures 1 and 4), yet this change was sufficient to make the figures discriminable by three- and four-year-olds under tachistoscopic conditions. The observation that an irrelevant part of a stimulus influences recognition in the young child is consistent with an earlier proposal (Ghent, 1960) that young children have difficulty using a part cue to identify the whole figure: the whole figure must be processed for identification to take place. Evidence that young children have difficulty in using a part cue has also been provided by Goldstein and Mackenbergh (1966).

EXPERIMENT V

This experiment extended the work on identification of geometric shapes to adults and asked whether recognition is influenced by the orientation in which the shapes are presented. The approach to the problem was essentially that used with the first-graders in Experiment II.

Method

The subjects were 32 male and 48 female college students, most of them undergraduate volunteers from an introductory class in psychology at George Washington University. Sixty-nine of them were right-handed and 11 were left-handed, the criterion being the hand used for writing.

The figures were the ten complex figures (Figure 1) presented to the first-graders in Experiment II. The figures were presented tachistoscopically in each of four orientations such that the distinguishing feature appeared at the top, the bottom, the left, or the right of the figure. The 40 items were arranged in a semirandom order in four different sequences. The subjects responded by selecting the item just seen from an array of 6 items of the appropriate family; the multiple-choice arrays were presented to the subject in the form of a booklet and he circled the figure he had just seen. As with the children, four sets of multiple-choice arrays were used, each set being oriented to correspond to one of the orientations of the stimulus figures (all items with the distinguishing feature at the top, all with the distinguishing feature at the bottom, all with the distinguishing feature at the left, and all with the distinguishing feature at the right). Thus the orientation of all the multiple-choice items was the same for any particular subject but differed among groups, whereas the stimulus figures appeared in each of the four orientations for all subjects regardless of group. Within each multiple-choice group, the subjects were counterbalanced with respect to the sequence of presentation used and to sex.

The subject sat before a table on which was placed a multiple-choice booklet. On another table straight ahead, 87 in. away, was a milk-glass screen 18 by 19 in. on a side, with a small, $2\frac{1}{4}$ -in., square outlined in fine black tape somewhat above the middle of the screen, approximately at the subject's eye level. The stimuli were projected from behind the screen by a Keystone projector fitted with an Ilex shutter. The image sizes averaged $1\frac{5}{8}$ in. for the long axes and subtended a retinal angle of approximately 1 deg. Ambient light was kept constant by drawing the window shades and leaving one of the overhead room lights on.

A set of 16 practice items always preceded the test items. The practice items, the same as those used with the six-year-olds, consisted of two families constructed so that the members of each family were alike on three sides and differed on the fourth side. The multiple-choice arrays for these items constituted the first 16 pages of the subject's response booklet, and from his point of view, this first set of trials did not differ from the next set except in the families of figures presented. Before the practice trials, he was told that geometric forms would appear in any of four orientations, in a random order, and that he was to try to find the correct figure in the multiple-choice array, disregarding orientation. After each trial, he turned the page in the booklet and circled his choice; he was not informed of the accuracy of his choice. He was told not to rotate the booklet and to proceed to the next page only after the stimulus had been presented. Before each trial, he was told to look at the square where the stimulus would appear. Except for some early practice trials, the practice figures and the test figures were presented at $\frac{1}{100}$ sec for all subjects.

Results and discussion

The main question is whether the orientation of the stimulus figures had any effect on recognition. For the vertically oriented figures, recognition was markedly better when the distinguishing feature appeared at

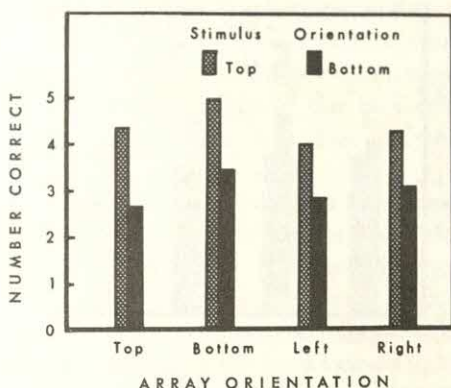


Fig. 6. The number of items identified correctly by college students when the distinguishing feature appeared at the top or at the bottom of each stimulus, for each array group receiving the multiple-choice items oriented differently, Experiment V

the top of the figure than when it appeared at the bottom. Figure 6 shows the mean number correct for the two orientations for each of the four groups receiving the multiple-choice arrays in a different orientation. An analysis of variance evaluated the effect of stimulus orientation (repeated measures) and array orientation; only the stimulus orientation exerted a significant effect ($F = 33.5$, $p < .01$ for $\frac{1}{2}$ df).

For the figures oriented horizontally, recognition was better when the distinguishing feature appeared at the left, as can be seen in Figure 7. Although a visual comparison of the scores for the groups shown the left and right arrays suggests that identity of orientation between stimulus and array figures facilitated correct identification, the analysis of variance did not yield a significant interaction between array and stimulus orientation; the only significant effect was that of stimulus orientation ($F = 7.31$, $p < .01$ for $\frac{1}{2}$ df).

The males and females showed the same patterning of scores with respect to the effects of orientation, but the various subgroupings of females usually showed higher total scores than the males. The mean total score for the females was 16.3 correct and that for the males was 13.1 ($t = 2.10$, $p < .05$). This result is somewhat surprising, since two previous studies—one with college students (Braine, 1965) and one with children (Ghent, 1960)—revealed significantly better performance by the males; but these latter studies used realistic figures with multiple distinguishing features.

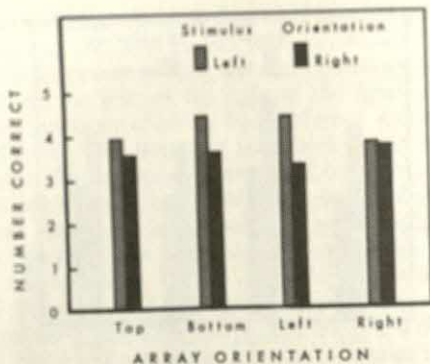


Fig. 7. The number of items identified correctly by college students when the distinguishing feature appeared at the left or at the right of each stimulus, for each array group receiving the multiple-choice items oriented differently, Experiment V

Previous work on the relation between recognition and stimulus orientation has focused on the question of whether an unfamiliar orientation impairs recognition of geometric shapes (Dearborn, 1899; Rock, 1956; Rock and Heimer, 1957) or of representational shapes such as pictures (Gibson and Robinson, 1935; Rock, 1956; Rock and Heimer, 1957) and letters (Kolars, 1968; Kolars and Perkins, 1969a; Kolars and Perkins, 1969b). It is clear from the data of Experiment V that recognition of geometric shapes by adults *can* be influenced by stimulus orientation in the absence of familiarity effects. In fact, it could be claimed that a familiar orientation was induced by the orientation of the figures in the multiple-choice array, yet an identity of orientation between stimulus and array figures did not facilitate recognition. The finding that orientation influences recognition in the absence of familiarity effects is consistent with the view that a serial-processing mechanism in form perception has a directional component.

GENERAL DISCUSSION

The three experiments using brief exposure durations of stimuli revealed that a major change occurs around four years of age in the role played by stimulus orientation in the recognition of geometric shapes. Three-year-olds recognized forms most readily when the forms were phenomenally upright, regardless of what part of the form contained the discriminandum, whereas five-year-olds and older subjects recog-

nized geometric forms best when the discriminandum was located at the top of the figure, regardless of the phenomenal orientation of the figure. The data also suggest that another change occurs, some time after seven years of age, in the discriminability of the two sides of horizontally oriented figures, but this change is less clear and was less well documented.

The discussion that follows will focus on four issues raised by these results. First, alternative interpretations for the older subjects' better recognition for the top of the figure will be considered, interpretations which involve the factors of attention and acuity. Second, the systematic studies of Rock and his colleagues on the relation between the orientation of a form and its recognition will be discussed in light of the present findings. Third, the developmental change in recognition of the distinguishing feature at the left and right sides of the figure will be related to results obtained by other investigators and to some cross-cultural material. Finally, the discussion will turn to the somewhat vexing problem of the definition of the 'focal' feature.

Alternative interpretations

It may be thought that the present findings in adults are not new, since many studies have found that items appearing in the upper part of an array are identified more often than the items appearing in the lower part (Harcum, 1964; Heron, 1957; Kimura, 1959; Takala, 1951).³ In general, the differential accuracy reported in those studies has been interpreted as due, in whole or in part, to response biases for dealing with an array of separate items. The present results cannot be accounted for in such terms. A single figure was presented, the figure was unfamiliar, and only one specific portion of the figure had to be perceived for identification of the whole figure. It seems clear that factors such as order of report, strategies for organizing or remembering a series of items, or habits developed for alphanumeric material cannot be invoked.

It might be claimed that there is some mysterious tendency to 'attend' to the top of a figure, but there is evidence against this interpretation by *attention*, at least in the simple meaning of the word. Both Sperling (1960) and Mewhort (1967) found better recognition for the top line of an array of letters, even when the method of partial report was used—a method which controlled order of report and attentional factors, since the subject was told to which line to direct his attention on each trial. Also, in an unpublished study by L. Humphrey in the author's laboratory, instructions to attend to the bottom of the figure did not reverse

accuracy scores for figures with the distinguishing features presented at top and at bottom. Invoking attention as something not characterized by voluntary control becomes problematical, because then it may not be possible to separate attentional and perceptual mechanisms. For example, if attention to the top of a figure is obligatory, then attention to the top would be indistinguishable from pattern processing that started at the top. Some of the contemporary views of attention (Neisser, 1966) explicitly propose that attention is the manifestation of a particular kind, or level, of pattern processing.

Another interpretation of the greater identifiability of the feature at the top of the pattern is that *acuity* is greater for the upper visual field than the lower. Since the stimuli used in this and other studies have appeared in the center of the field, the results could be due to the nature of pattern processing or to differential acuity. Although standard texts do not refer to any differences in acuity between the fields, Weymouth, Hines, Acres, Raaf, and Wheeler (1928) did report that acuity was better in the upper field for distinguishing the orientation of gratings exposed for .10 sec.⁴

Yet, the results of other studies in which stimuli have been presented separately to upper and lower fields do not present a consistent picture. Mishkin and Forgays (1952) reported that identification of words was significantly better in the lower visual field than the upper. Two subsequent studies did not replicate this finding (DeYoung, 1961; A. Sheldon, unpublished study done at George Washington University). A recent investigation of this problem with geometric shapes (Humphrey, 1970) yielded complicated results: recognition was better in the upper field, but this effect was related to the orientation of the multiple-choice array. The question of whether the upper and lower visual fields differ, either in the perceptibility of patterns under conditions of limited exposure or in acuity, or both, is not yet answered.

Several other considerations speak against interpreting the greater identifiability of the feature at the top as due primarily to an acuity difference. First, the presence or absence of a fixation point does not appear to affect the results obtained with geometric figures (unpublished work), which is surprising if small differences in acuity between the fields are indeed critical in producing differences in identifiability between the top and bottom parts of a figure. Second, the developmental data are difficult to account for in terms of acuity differences. Three-year-olds identified the figures better when the distinguishing feature appeared at the bottom of the pattern (presumably in the lower visual field), since this particular orientation was perceived as upright by the child. If differen-

tial recognition for figures in various orientations were the reflection of acuity differences, then one would have to conclude that acuity is better in the lower visual field in young children, and then changes abruptly to better acuity in the upper field between four and five years.

Since the alternative explanations of attentional or acuity differences do not seem satisfactory, it is concluded that the better recognition for the top part of a pattern (in subjects of five years and older) is due to a serial-processing mechanism. It might be thought that serial processing that goes from top to bottom derives from a reading habit, but the results speak strongly against this interpretation. First, the greater perceptibility for the top did not depend on the use of letters or numbers but appeared for a single unfamiliar geometric shape. Second, and more importantly, the effect appeared abruptly in children between four and five years, and was as strong in subjects of six years as in those of college age. It could be argued, as has been done in another context (Ghent, 1961), that the tendency to process from the top of the pattern itself provides the basis for the apparently universal tendency to read by starting at the top.

Studies of Rock et al.

The particular figures used in Experiments I-V permit comparison of the effect of two factors on recognition: the phenomenal uprightness of the figure and the position of the distinguishing feature. The data show that the figures used here were phenomenally upright for children when the family (or focal) feature was at the top; and similar judgments were obtained from a small group of college students (unpublished). Nevertheless, recognition by subjects older than four years was better when the distinguishing feature was at the top, which was the phenomenally upside-down orientation for the particular figures used here. This apparent irrelevance of the phenomenal upright conflicts with the view of Rock and his colleagues (Rock, 1956; Rock and Heimer, 1957; Rock and Leaman, 1963) that recognition in adults is impaired when the phenomenal top of the figure is not at the top of the phenomenal spatial framework. There are many reasons for the difference in view; perhaps the most important one is that the two sets of experiments were oriented toward different questions.

Rock, starting from the everyday experience that a change in orientation changes the appearance of the form, selected forms for which recognition was impaired when the figure was disoriented, and he asked whether that impairment was due to disorientation with respect to the retina or to the environment. The question of how, or why, disorien-

tation of the figure impairs recognition, regardless of which frame of reference was used, was not a primary concern.

By contrast, the burden of the argument here is that the effect of orientation on recognition is related to the specific characteristics of the figure; that is, some orientations place the relevant cue in locations that facilitate recognition, and other orientations do not. A somewhat similar point has been made by Kolers and Perkins (1969a, 1969b) on the recognition of English letters. This view also requires an answer to the question of what framework—retina or environment—defines the 'top' from which serial processing begins, but it further asks how particular characteristics of shapes interact with certain frameworks.

The other major difference in point of view between Rock's work and the present work is that Rock et al. seem to assume that a change in the phenomenal appearance of a form necessarily entails a change in the identifiability of the form. In their experiments, they asked whether a shape presented in one orientation is recognized, in general, when it is presented subsequently in a new orientation. Their subjects' responses could have been based on a judgment of phenomenal similarity ('this picture of a cat looks the same as it did before') or of identification ('this picture is a cat'). By contrast, the experiments above asked only for the identification of a shape presented in different orientations; for example, the subjects were specifically instructed to ignore the difference in phenomenal appearance between the stimulus and multiple-choice items, and to pick the item of the same form as that just seen. This assumes that some changes in phenomenal appearance of a form may occur without concomitant changes in identifiability. The issue is discussed in detail in another paper (Braine, in preparation), where it is argued that a change in the orientation of a figure is always associated with a change in the phenomenal appearance of the figure for the adult but that whether or not identifiability is affected depends on the nature and location of cues available for identifying the figure.

In addition to differences in the questions asked, there were many procedural differences between the two sets of experiments. In most of the experiments by Rock et al. their subjects had no information on the orientation of the stimuli to be presented and likely implicitly assumed that the pictures would be presented in their usual orientations. That incorrect assumption could produce the impairment in recognition found when the top of the stimulus does not coincide with the phenomenal top of the framework (see Braine, 1965). Another ambiguity in the task occurred when standard stimuli were presented in one orientation but comparison stimuli in two orientations, allowing the subjects the option

of rejecting a comparison figure that was recognized as the same shape as the standard but as different in orientation. Although Rock acknowledged this problem, the procedure subsequently employed (Rock and Heimer, 1957) did not appear to eliminate the ambiguity of whether the response reflected recognition of a form or recognition of a form in a specific orientation.

Horizontally oriented figures

With the long axis of the figure horizontal, college students identified the figures better when the distinguishing feature appeared on the left side of the figure than on the right. This finding may not seem new, since it does have a parallel in two other sets of work: the studies of Takala (1951), and those of American investigators such as Bryden (1966), Harcum (1964), Heron (1957), and Kimura (1959). Takala conducted a most thorough and inventive analysis of anisotropies of visual space, and concluded that the upper and left portions of space are favored with respect to the memory for items in these portions of the field, rotations of figures in reproduction, judgment of movement, their role in ambiguous figures, and so on. American investigators have reported better identifiability for items located in the left part of an array of items, usually letters.

However, the point made earlier for the vertically oriented figures must be made again: differential recognition found in studies using an array of items and involving the report of multiple items is probably due to an unknown combination of perceptual and response factors. The unique aspect of the finding reported in this paper is that differential recognition appeared even when a single, geometric, unfamiliar figure was used; it is likely that perceptual discrimination itself was better for the left side of the figure. In analogy with the line of thought developed for the vertically oriented figures, the better recognition of the left would be due to serial processing beginning with that side. Such a tendency would seem very natural to readers or writers of English and could be attributed to a reading habit. However, this point has been investigated directly by testing recognition in Israelis whose native reading language proceeds from right to left, with the surprising finding that better recognition appeared on the left in these college students too (Braine, 1968). The situation is further complicated by the findings on children.

For example, American first-graders showed a slight tendency to discriminate the distinguishing feature more often when it appeared on the right (Experiment II); the same effect was shown by Israeli children from elementary school (Braine, 1968); and Dyer and Harcum (1961) re-

ported that first-graders identified the right side of binary patterns better than the left side. These developmental changes could be accounted for by assuming that there is a tendency to process horizontally oriented stimuli in a rightward direction and that the starting position changes with age (Braine, 1968). For the adult, the starting position would be at the extreme left of the pattern, whereas for the children, the starting position would be the center of the figure or wherever fixation happened to be at the time of the exposure. The view is very similar to that proposed to account for the developmental changes observed for the vertical axis, although the ages of transition are very different. The analogy would be complete if it could be assumed that a focal feature determined the starting position for the horizontal axis in children, but there does not seem to be any evidence for this assumption.

The superficial similarity of the developmental data for vertically and horizontally oriented figures might permit a common interpretation for both cases. But despite the attraction of this parsimonious approach, and my earlier espousal of it (Braine, 1968), several aspects of the data make it unlikely that the same processes underlie the two lines of development. First, the difference in recognition for the top and bottom is a much more robust effect than is the difference for the left and right; a similar conclusion has been drawn by Takala (1951) on the basis of a much wider range of stimulus materials. Secondly, the age at which the starting position is defined by the spatial framework is very different for vertically and horizontally oriented figures. Finally, the focal feature of a horizontally oriented figure does not appear to influence either recognition or judgments of the phenomenally 'right' orientation of a horizontal figure. There would seem to be no reason for these differences between horizontally and vertically oriented figures were the serial-processing mechanisms the same in both cases (except for a difference in the directionality of the processing). Although an alternative model to account for the results with horizontally oriented figures would be desirable, none seems available.

Definition of the 'focal' feature

The term 'focal' plays a central role in the theory of the serial processing of form by very young children. The concept first arose in an attempt to explain the observation that young children consider some geometric shapes upright in one orientation and upside down when turned 180 deg. The only explanation that appeared to account for these earlier data was that the child judged a form to be upright when the

feature that appeared to be salient was at the top of the figure or the card.

The alternative explanation most commonly offered is that forms are judged upright when in a position of maximum stability, and this interpretation might indeed account for the judgments of the distinguishing-feature figures described in the present paper. However, the notion of 'stability'—not easily definable itself—could not also encompass the judgments of cards described in earlier work (Ghent, 1961).

First, most of the cards in the 1961 study were such that there was no *a priori* basis for considering one orientation more stable than the other. For example, circular figures were judged upright significantly often in one orientation, although the figure would be expected to 'fall over' in both directions. Similarly, cards containing a dot or a line were considered upright with the figure in the upper part of the card rather than the lower; a single vertical line was considered upright when the line was differentiated (by another line or gap) in the upper part rather than the lower part; and a card was considered upright when the card as a whole was differentiated in the upper rather than the lower part (see Ghent, 1961, Figure 2). Yet it is difficult to see why differentiation in the upper part of the card should favor stability.

Second, there were some judgments of uprightness clearly contrary to those predicted on the basis of stability. At least four cards were of this kind (Ghent, 1961, Figure 1, items 3, 7, 9; Figure 2, item 10), and several others could be included in this category if the assumption were made that a black portion is perceived as heavier than a white portion. In addition, there were four pairs of figures in which one member of each pair was unequivocally more stable than the other member (since one was a variation on a square and the other was the identical variation on a circle), yet the responses were similar to both members of the pairs.

One of the main justifications for the use of the term 'focal' is that it can be defined in a principled (nonintuitive and noncircular) manner for very simple stimuli: the focal feature is the one differentiating feature in an otherwise homogeneous figure or card. Stimuli designed on the basis of this *a priori* definition of focal were indeed judged to be upright when this focal feature appeared at the top (Ghent, 1961, Figure 2), an outcome which, as mentioned above, could not be predicted on the basis of stability. Also, the judgments were not unique to American children but were replicated with Iranian preschool children (Antonovsky and Ghent, 1964). By extension, the focal feature for complex figures can be defined functionally—behaviorally—as whatever kind of feature the young child prefers at the top.

On the basis of this functional definition, then, a partial enumeration of focal features is possible for figures symmetrical about the vertical axis: convex curves pointing upward, in contrast to emptiness or straight lines; angles pointing upward, in contrast to emptiness or straight lines; straight horizontal lines, in contrast to emptiness; and patterned portions, in contrast to plain portions. Several other contrasts could be made, since the definition of the focal feature is, in part, a relative matter, but considerably more data are required in order to establish a reliable hierarchical ordering. It is tempting to try to define the focal feature as the part that is most complex, complexity being defined either by the number of turns or by the amount of contour. However, some of the figures reported in this paper indicate that the more complex portion is not necessarily the focal feature (Figure 1, last two items in third family; Figure 4, last item in each family).

In addition to the fact that 'focal' can be defined for at least simple stimuli, the concept provides a basis for interpreting the preferences for orientation in the context of a theory of form perception rather than as aberrations of childhood. For the young child, a figure is upright when the orientation conforms to, or facilitates, serial processing of the parts in the preferred order, namely, downward from the focal feature. A figure is upside down when the orientation opposes, or is incongruent with, the preferred order of processing. This interpretation led to the experiments on recognition reported here. If the studies on recognition were to be considered without reference to the judgments of orientation, it would be possible to dispense with the term 'focal,'⁶ but there would be considerable loss in the range of data accommodated.

In conclusion, this paper has presented a developmental analysis of the effect of stimulus orientation on the recognition of a certain type of geometric shape. The point of view developed assumes that a serial-processing mechanism operates in form perception, and proposes that this processing is directional rather than random. Directionality is divided into two aspects: a starting position, and a direction of subsequent processing of the parts. For figures in which the main axis is vertical, processing proceeds in a downward direction, and the starting position is the focal feature for the very young child and the top of the figure for the older child and adult. This line of thought has implications for the way in which orientation may influence the phenomenal appearance, and the identifiability, of a form; these implications will be discussed in another paper.

Notes

This work was supported by Grants MH04109 and HD05035 from the National Institutes of Health. The author thanks Walter Gogel for his helpful comments on an earlier version of this paper; Ronald Loughrey and Renee Chambliss for their help in testing the children in Experiments II and IV; and Evans Mandes, who tested most of the subjects in Experiment V. She also thanks Dr. B. Hyams and the teachers of the University of Maryland Nursery School; Dr. R. R. Foster of the Department of Research, the principal and teachers of the Grant School in Washington, D.C.; Mr. P. Weimer and the teachers of the Piedmont Nursery School for their patience and cooperation during a long period of testing. Received for publication April, 1971.

1. A brief report of the tachistoscopic results of Experiment IV appeared in *Psychonomic Science*, 1965, 2:155-156.

2. I am much indebted to Mrs. Robert Headley, who undertook the investigation of this problem.

3. Better recognition for the upper part of a display or of the visual field also seems to have been found by Wolford, Wessel, and Estes (1968), but these authors considered their results specific to their displays.

4. Thanks to Mrs. Linda Humphrey for bringing this reference to my attention and for providing many stimulating discussions on the problem of differential acuity in the two fields.

5. If future work should indicate that patterns presented separately to upper and lower fields are identified more readily in the upper field, that finding could be incorporated into the line of thought already developed by redefining the 'top' for serial processing as the top of the visual *field* rather than the top of a particular *figure* in the field. Then serial processing would always start in the upper field, and the trace of a figure presented in the lower field would be processed later in the sequence than a trace in the upper field.

6. The developmental change would be interpreted in the same way as that proposed for horizontally oriented figures. Downward processing for the young child would start at the center of the figure or wherever fixation happened to be at the time of exposure, whereas downward processing for the older child and adult would start at the top of the pattern.

References

- Antonovsky, H. F., and Ghent, L. 1964. Cross-cultural consistency of children's preferences for orientation of form. *American Journal of Psychology* 77:295-297.
- Braine, L. G. 1965. Disorientation of forms: An examination of Rock's theory. *Psychonomic Science* 3:541-542.
- Braine, L. G. 1968. Asymmetries of pattern perception observed in Israelis. *Neuropsychologia* 6:73-88.
- Bryden, M. P. 1966. Accuracy and order of report in tachistoscopic recognition. *Canadian Journal of Psychology* 20:262-272.
- Dearborn, G. V. N. 1899. Recognition under objective reversal. *Psychological Review* 6:395-406.

- De Young, K. N. 1961. Word recognition in the upper and lower visual fields. Unpublished doctoral dissertation, University of Minnesota. *University Microfilms* 61:5,838.
- Dyer, D. W., and Harcum, E. R. 1961. Visual perception of binary patterns by preschool children and by school children. *Journal of Educational Psychology* 52:161-165.
- Ghent, L. 1960. Recognition by children of realistic figures presented in various orientations. *Canadian Journal of Psychology* 14:249-256.
- Ghent, L. 1961. Form and its orientation: A child's-eye view. *American Journal of Psychology* 74:177-190.
- Ghent, L. 1964. The effect of orientation on the recognition of geometric forms by retarded children. *Child Development* 35:1127-1136.
- Ghent, L., and Bernstein, L. 1961. Influence of the orientation of geometric forms on their recognition by children. *Perceptual and Motor Skills* 12:95-101.
- Gibson, J. J., and Robinson, D. 1935. Orientation in visual perception: Recognition of familiar plane forms in different orientations. *Psychological Monographs* 46(no. 210):39-47.
- Goldstein, A. G., and Mackenberg, E. J. 1966. Recognition of human faces from isolated facial features: A developmental study. *Psychonomic Science* 6:149-150.
- Harcum, E. R. 1964. Reproduction of linear visual patterns tachistoscopically exposed in various orientations. College of William and Mary.
- Harris, L., and Schaller, M. J. 1969. Form and its orientation: Reexamination of a child's-eye view. Paper presented at the meeting of the Psychonomic Society, Saint Louis.
- Hebb, D. O. 1949. *The organization of behavior*. New York: Wiley.
- Howard, I. P., and Templeton, W. B. 1966. *Human spatial orientation*. New York: Wiley.
- Heron, W. 1957. Perception as a function of retinal locus and attention. *American Journal of Psychology* 70:38-48.
- Humphrey, L. 1970. Form recognition in upper and lower visual fields. Unpublished master's thesis. University of California at Santa Barbara.
- Kimura, D. 1959. The effect of letter position on recognition. *Canadian Journal of Psychology* 13:1-10.
- Koff, E. B. 1969. Preference for orientation of nonrealistic forms. Paper presented at the meeting of the Eastern Psychological Association.
- Kolers, P. A. 1968. The recognition of geometrically transformed text. *Perception and Psychophysics* 3:57-64.
- Kolers, P. A., and Perkins, D. N. 1969a. Orientation of letters and errors in their recognition. *Perception and Psychophysics* 5:265-269.
- Kolers, P. A., and Perkins, D. N. 1969b. Orientation of letters and their speed of recognition. *Perception and Psychophysics* 5:275-280.
- Mayzner, M. E., Tresselt, M. E., Tabenkin, N., Didner, R., and Helfer, M. S. 1969. Visual information processing of sequentially presented inputs: I. Effects of input timing on sub-span storage and retrieval mechanisms. *Perception and Psychophysics* 5:297-302.
- Mewhort, D. J. K. 1967. Familiarity of letter sequences, response uncertainty,

- and the tachistoscopic recognition experiment. *Canadian Journal of Psychology* 21:309-321.
- Mishkin, M., and Forgy, D. G. 1952. Word recognition as a function of retinal locus. *Journal of Experimental Psychology* 43:43-48.
- Moeller, C. E., and Goodnow, J. J. 1969. Orientations called "right-side-up": Effects of stimulus alignment. *Psychonomic Science* 16:213-215.
- Neisser, U. 1966. *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Pick, H. L., Klein, R. E., and Pick, A. D. 1966. Visual and tactual identification of form orientation. *Journal of Experimental Child Psychology* 4: 391-397.
- Rock, I. 1956. The orientation of forms on the retina and in the environment. *American Journal of Psychology* 69:513-528.
- Rock, I., and Heimer, W. 1957. The effect of retinal and phenomenal orientation on the perception of form. *American Journal of Psychology* 70:493-511.
- Rock, I., and Leaman, R. 1963. An experimental analysis of visual symmetry. *Acta Psychologica* 21:171-183.
- Sperling, G. 1960. The information available in brief visual presentations. *Psychological Monographs* 74(no. 11, whole no. 498).
- Takala, N. 1951. *Asymmetries of the visual space*. Helsinki: Suomalaisen Kirjallisuuden Seuran Kirjapainon Oy.
- Weymouth, F. W., Hines, D. C., Acres, L. H., Raaf, A. E., and Wheeler, M. C. 1928. Visual acuity within the area centralis and its relation to eye movements and fixation. *American Journal of Ophthalmology* 11:947-960.
- Wolford, G. L., Wessel, D. L., and Estes, W. K. 1968. Further evidence concerning scanning and sampling assumptions of visual detection models. *Perception and Psychophysics* 3:439-444.

continued from page 309

- Chess, Stella, and Thomas, Alexander, eds. *Annual Progress in Child Psychiatry and Child Development*. New York: Brunner/Mazel, 1971. Pp. x, 614. \$15.00.
- Christenseon, Cornelia V. *Kinsey: A Biography*. Bloomington: Indiana University Press, 1971. Pp. xii, 241. \$6.95.
- Cole, Michael; Gay, John; Glick, Joseph A; and Sharp, Donald W. *The Cultural Context of Learning and Thinking*. New York: Basic Books, 1971. Pp. xx, 304. \$10.00.
- Doob, Leonard W. *Patterning of Time*. New Haven, Conn.: Yale University Press, 1971. Pp. xiv, 472. \$15.00.
- Elbert, Edmund J. *I Understand: A Handbook for Counselling in the Seventies*. New York: Sheed and Ward, 1971. Pp. ix, 291. \$6.95.
- Eleftheriou, B. E., and Scott, J. P., eds. *The Physiology of Aggression and Defeat*. New York: Plenum Publishing, 1971. Pp. xi, 312. \$13.50.
- Elliott, C. K. *A Guide to the Documentation of Psychology*. Hamden, Conn.: Shoe String Press, 1971. Pp. 134. \$5.50.
- Finch, Stuart M., and Poznanski, Elva O. *Adolescent Suicide*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xv, 66. \$6.75.
- Fox, Michael W. *Integrative Development of Brain and Behavior in the Dog*. Chicago: University of Chicago Press, 1971. Pp. xii, 348. \$22.50.
- Glidewell, John C. *Choice Points: Essays on the Emotional Problems of Living with People*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. 144. Paperback, \$1.95.
- Grosser, George H.; Wechsler, Henry; and Greenblatt, Milton, eds. *The Threat of Impending Disaster: Contributions to the Psychology of Stress*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xi, 335. Paperback, \$2.95.
- Hale, Nathan G., Jr., ed. *James Jackson Putnam and Psychoanalysis*. Cambridge, Mass.: Harvard University Press, 1971. Pp. xiv, 384. \$14.00.
- Hand, Wayland D., ed. *American Folk Legend: A Symposium*. Berkeley and Los Angeles: University of California Press, 1971. Pp. vi, 237. \$7.50.
- Harth, Robert, ed. *Issues in Behavior Disorders: A Book of Readings*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xviii, 674. \$26.75.
- Hill, John P., ed. *Minnesota Symposia on Child Psychology*. Vol. 5. Minneapolis: University of Minnesota Press, 1971. Pp. xi, 216. \$5.00.
- Homme, Lloyd; Csanyi, Attila P.; Gonzales, Mary Ann; and Rechs, James R. *How to Use Contingency Contracting in the Classroom*. Champaign, Ill.: Research Press, 1971. Pp. xi, 130.
- Howells, John G., ed. *Modern Perspectives in International Child Psychiatry*. New York: Brunner/Mazel, 1971. Pp. xviii, 878. \$25.00.
- Hursch, L. *Der Einfluss verschiedener Versuchssituationen auf die Faktorenstruktur von Intelligenzleistungen*. [With a summary in English.] Berne: Hans Huber, 1970.
- Isaacson, Robert L., and Hutt, Max L. *Psychology: The Science of Behavior*. 2d ed. New York: Harper & Row, 1971. Pp. x, 433.
- Jacobs, Alfred, and Sachs, Lewis B., eds. *The Psychology of Private Events*. New York: Academic Press, 1971. Pp. vi, 201. \$10.50.
- Jarrard, Leonard E., ed. *Cognitive Processes of Nonhuman Primates*. New York: Academic Press, 1971. Pp. xii, 188. \$9.50.

continued on page 236

EXAMINATIONS OF THE CATEGORY-RECALL FUNCTION

Robert M. Schwartz and Michael S. Humphreys

University of British Columbia

With experimentally paced presentation, subjects allowed to choose their number of categories showed a category-recall function of less magnitude than has been reported with self-paced tasks, and subjects assigned their number of categories showed little evidence of the function. Individual differences in categorization accounted for some of the variance, but individual differences in recall ability accounted for more.

Tulving (1962) found that in multitrial free recall, increases in organization over trials parallel increases in recall performance and that subjects who show high organization recall more than subjects who show low organization. These findings led Tulving to conclude that organization is an important determinant of recall performance.

However, because Tulving's findings were correlational, it is possible that there is a third variable which affects both organization and recall performance. Earhard suggested that organization may reflect some individual-differences variable, and he supported this suggestion by showing that when organization was measured independently on a free-recall task, those organization scores predicted performance on both paired-associate and serial-learning tasks (Earhard, 1967; Earhard and Endicott, 1969).

Mandler (1967) used a paradigm which, on the surface, seems to bypass the correlational nature of Tulving's original findings by providing a measure of organization before asking subjects to recall. Typically here, subjects sort words into categories; they are usually allowed to choose the number of categories they use, with the restriction that they use no less than two and no more than seven categories. The sorting process is continued until stable categorization is reached. Then, they are asked to recall as many words as they can. Mandler generally found a linear relationship between the number of categories used and the amount recalled—the category-recall function. The relationship was usually characterized by a slope value of 5 ± 2 , which indicates that subjects recalled an additional 5 ± 2 words per category used.

Mandler (1968) identified two possible artifacts of his experimental paradigm. First, as sorting is self-paced and the subject sorts to a criterion, the number of sorting trials and the time per sorting trial are not experimentally controlled. Second, allowing each subject to choose his own number of categories may result in an artifact of self-selection. Mandler used variations of his paradigm to investigate these possible artifacts. He found that the category-recall function held when subjects sorted for a fixed number of trials and were presented words at a fixed rate (Mandler, 1967, Experiment H) and when they were assigned their number of categories in either a between-subjects (Mandler, 1967, Experiment B) or a within-subjects (Mandler, 1968) design. However, in the experiment controlling number of trials and time per trial, the subjects were allowed to choose their number of categories; and in the experiments in which they were *assigned* their number of categories, number of trials and time per trial were not experimentally controlled.

In the present experiments the category-recall function was examined for three purposes: (a) to replicate Mandler's results when time per trial and number of trials were experimentally controlled and subjects were allowed to choose their number of categories, (b) to extend the category-recall function to the case where time and trials are experimentally controlled and subjects are assigned their number of categories, and (c) to determine the relationship between the number of categories a subject chooses and his ability to recall.

EXPERIMENT I

The first step in determining whether the number of categories chosen reflects individual differences in recall ability was to determine whether subjects habitually sort into a particular number. In Experiment I each subject sorted five lists of words, and the differences between subjects in mean number of categories were examined.

Method

Stimuli were 260 nouns randomly chosen from the Paivio, Yuille, and Madigan (1968) norms. Each noun was typed on a 3-by-5-in. card, and the cards were divided into five lists of 52 words each. Each subject was given the five lists successively and was told to sort the words in a given list into between two and seven categories, using any kind of sorting strategy he desired. Each subject was given one sorting trial on each list. The subjects were ten student employees of the psychology department.

Results

Each subject was given five scores indicating the number of categories he used on each of his five sorts, and an average of these scores was calculated for each subject. The range of these averages was 2.2 to 7.0, and the median was 6.0. The data were subjected to analysis of variance, and the results indicate that there was a significant difference between subjects in the number of categories [$F(9, 40) = 7.76, p < .001$]. The significant value indicates that there were consistent individual differences in the number of categories the subjects used in this word-sorting task.

EXPERIMENT II

In Experiment II the subjects sorted two lists, both presented for a fixed number of trials and at a fixed rate. On the first list, the subjects were allowed to choose their number of categories, and thus, the first part of the experiment was a replication of Mandler's (1967, Experiment H). On the second list, the subjects were assigned their number of categories.

Two analyses of second-list performance were used to examine individual differences. First, since Experiment I indicated that there were individual differences in choice of number of categories, second-list performance was examined as a function of a given subject's first-list choice of number of categories. Second, recall performance on the second list was analyzed as a function of his first-list recall performance.

Method

Two of the 52-word lists used in Experiment I served as stimuli. Five randomizations of the first list and four randomizations of the second list were tape-recorded for presentation; the interval between words of a given randomization was 4 sec, and the interval between randomizations of a given list was 10 sec.

The subjects were told that they would be presented with 52 nouns, one at a time, and that their task was to put these nouns into categories. They were given answer booklets with columns and were told to let each column represent a category. The subjects were informed that they could use "anywhere between two and seven categories" and that on each trial they should try to use the same organization as on the previous trial. They were also told that they would be asked to recall the words at the end of the fifth sorting trial, and they were allowed 3 min for recall.

The procedure for the second task was the same as for the first task except

that each subject was randomly assigned either three or six categories and only four presentation trials were used.

With second-list recall as the dependent measure and number of categories on the first list as the measure of individual differences, the experiment was treated as a 2×2 fixed-factor design. A subject was designated a *high categorizer* if he used five, six, or seven categories on the first task and was designated a *low categorizer* if he used two, three, or four categories on the first task. High categorizers could be assigned either six (group HC-6) or three (group HC-3) categories in the second task, and low categorizers could be assigned either six (group LC-6) or three (group LC-3) categories in the second task.

With second-list performance as the dependent measure and first-list performance as the measure of individual differences, four groups again resulted. A subject who recalled above the median on the first list was designated a *high recaller*, and one who recalled below the median was designated a *low recaller*. High recallers could be assigned either six (group HR-6) or three (group HR-3) categories on the second list, as could low recallers (group LR-6 and group LR-3).

The subjects were 40 volunteers from introductory psychology classes at the University of British Columbia. They were run in three group sessions. Because both conceptualizations of the design were dependent on first-list results and a random assignment of subjects to second-list number of categories, unequal numbers of subjects appeared in each group. The numbers of subjects in each group when first-list number of categories was the measure of individual differences are contained in Table 1; the numbers of subjects in each group when first-list recall was the measure of individual differences are shown in Table 2.

Results

The subjects were divided (as described above) into groups of *high and low categorizers* on the basis of their first-list choice of number of categories. The mean number of categories used by groups HC-6 and HC-3 was 6.09, and the mean number of categories used by groups LC-6 and LC-3 was 3.56. Mean first-list recall scores for these high- and low-categorizer groups are shown in Table 1. These data were analyzed in order to determine (a) whether the category-recall function Mandler reported was replicated and (b) whether the assignment of subjects to three and six categories for second-list sorting showed any systematic bias in terms of first-list recall performance. Analysis of variance found no evidence for high categorizers to recall more than low categorizers [$F(1, 36) = .59$], indicating that the category-recall function was not replicated. Also, no evidence was found for sampling bias in second-list assignment, either as a main effect of number of second-list categories [$F(1, 36) = .94$] or as an interaction between number of first-list and number of second-list categories [$F(1, 36) = 1.11$].

Table 1. Summary data for high and low categorizers, Experiment II

	Group HC-3	Group HC-6	Group LC-3	Group LC-6
Number of subjects	11	11	12	6
Mean first-list recall	25.73	28.18	26.17	25.50
Mean second-list recall	21.09	26.81	23.92	19.50

The failure to find a significant difference between high and low categorizers and, thus, the failure to replicate the category-recall function could have been due to a loss of statistical power caused by dichotomizing the number-of-categories variable. To examine this possibility, a correlation between first-list recall and number of first-list categories was computed. The obtained coefficient was not significant [$r = .22$, $p > .10$]. Nevertheless, for comparison with Mandler's results, the line of best fit was also obtained. The resultant equation, $.6NC + 23.5$, where NC is number of categories, indicated that subjects added an average of .6 words per category used. The relationship between number of categories and recall was in the direction Mandler (1967) predicted, but the slope value of .6 was significantly less than the minimal slope of 3.0 he predicted [$t(38) = 5.48$, $p < .001$; see Draper and Smith, 1966, p. 19].

One reason why the category-recall function was not replicated could be that our subjects were not sorting as consistently as Mandler's (1967, Experiment H). The only statistic that Mandler reported on sorting consistency was that 3 of 42 subjects sorted identically on the last two trials; however, 2 of them did not sort by content. In the present study, 2 of 40 subjects sorted identically on the last two trials, and both used content sorts. Thus, the present data on sorting consistency compare favorably with Mandler's.

The data on number of subjects using identical sorts on the last two trials do little to describe the sorting consistency of the entire group. For this purpose, the percentage of items sorted identically on the last two trials was determined for each subject. The median of these percentages, 85%, indicates that the present subjects were sorting fairly consistently.

The mean *second-list* recall scores for the high and low categorizers are also shown in Table 1. Although these scores were slightly lower than the first-list recall scores, the difference is attributable to the use of five presentation trials on the first list and only four presentation trials on the second list. While the second-list recall scores were lower than the first-list recall scores, the subjects were not sorting less consistently on the second list. Three subjects sorted the second list identically on the

last two trials, and the median percentage of items sorted identically on the last two trials was 82%.

Analysis of variance was used to examine those second-list recall scores. The results showed no difference between high and low categorizers [$F(1, 36) = .96$] and no difference due to category assignment in the second task [$F(1, 36) = 1.16$]. However, the interaction of groups (high versus low categorizers) and category assignment was significant [$F(1, 36) = 9.87, p < .005$]. To control for possible artifacts due to assignment to groups, an analysis of covariance, holding first-list recall constant, was used on the second-list recall data. No difference between high and low categorizers was found [$F(1, 35) = .41$], nor any difference due to category assignment [$F(1, 35) = .38$]. However, the interaction of groups (high versus low categorizers) and category assignment was again significant [$F(1, 35) = 8.92, p < .01$]. Thus, as Table 1 shows, there appears to have been some tendency for subjects to recall more when they were assigned a number of categories that was similar to the number of categories they habitually use.

In order to analyze second-list performance as a function of recall ability, the subjects were also divided (as described in the section on method) into groups of *high* and *low* recallers according to their first-list recall performance. This division was at the median of first-list recall, and number of first-list categories was ignored. Subjects who recalled 27 or more words on the first list were assigned to the groups of high recallers, and those who recalled 26 or fewer words on the first list were assigned to the groups of low recallers. The mean first-list recall scores for those groups are shown in Table 2.

Table 2. Summary data for high and low recallers, Experiment II

	Group HR-3	Group HR-6	Group LR-3	Group LR-6
Number of subjects	13	9	10	8
Mean first-list recall	29.31	29.55	21.60	24.63
Mean second-list recall	25.00	25.89	19.40	22.38

The mean second-list recall scores for the (first-list) high and low recallers are also shown in Table 2. Analysis of variance was used to examine these scores. The group of first-list high recallers did significantly better on the second list than did the group of first-list low recallers [$F(1, 36) = 8.86, p < .01$]. The effect of category assignment was again not significant [$F(1, 36) = 1.13$], nor was the interaction of groups and category assignment [$F(1, 36) = .68$]. Thus, while the num-

ber of categories subjects were assigned had little or no effect on their second-list performance, constant individual differences, as measured by their first-list performance, had a major effect.

EXPERIMENT III

The failure to replicate Mandler's apparently well established category-recall function during learning of the first list warrants further investigation. Experiment III was thus conducted as a replication of Experiment II.

Method

The design, stimuli, and procedure were identical to those used in Experiment II. The subjects were 43 students enrolled in two sections of introductory psychology at Capilano College, a two-year community college. They were run during their regularly scheduled lecture hours. The numbers of subjects designated high and low categorizers are shown in Table 3, the numbers designated high and low recallers in Table 4.

Results

As in Experiment II, the subjects were divided into groups of *high and low categorizers* on the basis of their first-list choice of number of categories. The mean number of categories used by the high categorizers was 6.17, and the mean number of categories used by the low categorizers was 3.68. The mean first-list recall scores for those groups are shown in Table 3. Analysis of variance revealed that the high categorizers recalled more than the low categorizers [$F(1, 39) = 8.55, p < .01$], indicating that there was a significant category-recall function. No main effect of second-list category assignment was found [$F(1, 39) = 1.29$], but there was some evidence for an interaction of groups (high versus low categorizers) and category assignment [$F(1, 39) = 4.77, p < .05$], indicating that the random assignment of subjects to categories resulted in slightly better recallers being assigned to groups HC-6 and LC-3 than HC-3 and LC-6.

Table 3. Summary data for high and low categorizers, Experiment III

	Group HC-3	Group HC-6	Group LC-3	Group LC-6
Number of subjects	12	12	10	9
Mean first-list recall	29.08	30.17	27.70	22.00
Mean second-list recall	27.33	30.08	25.10	24.33

The correlation between number of first-list categories and first-list recall was significant [$r = .42, p < .01$]. The line of best fit, $1.7NC + 19.0$, where NC is number of categories, indicates that the subjects added an average of 1.7 words per category used. While this is a stronger relationship between number of categories and recall than was observed in Experiment II, the observed slope of 1.7 was still significantly less [$t(41) = 2.28, p < .05$] than the minimal slope of 3.0 predicted by Mandler (1967).

Although the subjects in Experiment III produced a significant category-recall function, they were not more consistent in sorting than the subjects in Experiment II. Two of them used identical sorts on the last two trials, and the median percentage of items sorted identically on the last two trials was 75%, which is 10% less than the corresponding value for Experiment II.

The mean *second-list* recall scores for the high and low categorizers are also shown in Table 3. As in Experiment II, second-list recall scores were slightly lower than first-list recall scores, since subjects were given one more presentation trial on the first list. As in Experiment II, the subjects were not less consistent in sorting the second list than they were in sorting the first list. Three sorted the entire list identically on the last two trials, and the median percentage of items sorted identically on the last two trials was 77%.

Analysis of variance was used to examine these second-list recall scores. The results indicate a tendency for high categorizers to recall more than low categorizers [$F(1, 39) = 3.18, p < .10$]. No effect of category assignment was found [$F(1, 39) = .35$], nor was there an interaction of groups (high versus low categorizers) and category assignment [$F(1, 39) = .57$].

Analysis of covariance, holding first-list recall constant, was also used to examine second-list performance of the high and low categorizers. This analysis eliminated the differences between the high and low categorizers [$F(1, 38) = .18$]. However, the effect of category assignment now approached significance [$F(1, 38) = 3.93, p < .10$], indicating a tendency for subjects assigned to use six categories to recall more than subjects assigned to use three categories when differences in first-list recall were statistically eliminated. The interaction of groups (high versus low categorizers) and category assignment was still not significant [$F(1, 38) = 1.09$], indicating that the bias in assigning subjects to conditions had not obscured that interaction.

In order to analyze second-list performance as a function of recall ability, the subjects were also divided, as in Experiment II, into groups

of high and low recallers, on the basis of their first-list scores. Subjects who recalled 28 or more words on the first list were assigned to the groups of high recallers, and those who recalled 27 or fewer words on the first list were assigned to the groups of low recallers. Mean first-list recall scores for these groups are shown in Table 4.

Table 4. Summary data for high and low recallers, Experiment III

	Group HR-3	Group HR-6	Group LR-3	Group LR-6
Number of subjects	12	9	10	12
Mean first-list recall	32.50	32.00	23.60	22.67
Mean second-list recall	31.00	31.11	20.70	25.00

Mean second-list recall scores for these groups of (first-list) high and low recallers are also shown in Table 4. Analysis of variance was used to examine these scores. The (first-list) high recallers did significantly better on the second list than did the (first-list) low recallers [$F(1, 39) = 18.12, p < .001$]. There was no effect of category assignment [$F(1, 39) = .48$], and the interaction of groups (high versus low recallers) and category assignment was not significant [$F(1, 39) = 2.18, p > .10$]. Thus, neither high nor low recallers recalled significantly more when assigned six as opposed to three categories.

DISCUSSION

Self-selected categories

For simplicity of interpretation, the results will be discussed in terms of the three purposes of the experiments. The first concerns the data on self-selected categories.

Both Experiments II and III resulted in a positive correlation between number of categories and recall. Although the correlation was significant only in Experiment III, the positive correlations in both studies, in conjunction with Mandler's (1967, Experiment H) findings of positive correlations, can be taken as supportive of the general form of the category-recall function when time and trials are experimentally controlled.

But even if it is concluded that there is a positive relationship between number of categories and recall, the present results question the magnitude of that function. Mandler (1967, Experiment H) reported correlations of .64 and .53 for two groups of subjects, while the present experiments found correlations of only .22 and .42. Nelson, McRae, and

Sturges (1970), in a study designed to replicate Mandler's, reported non-significant correlations of .34 and .15. Taken as a whole, these correlations for experimentally paced tasks are much lower than the median correlation of .70 that Mandler reported finding when sorting was self-paced and continued to a criterion.

The slopes relating number of categories to recall in Experiments II and III were .6 and 1.7 respectively; these slopes indicate that subjects were adding .6 and 1.7 words per category used. In his experiment with experimentally paced sorting, Mandler reported slopes of 3.0 and 3.7; in his experiments with self-paced sorting, Mandler reported a median slope of 3.9. Mandler argued that his slopes supported his assumption that subjects add 5 ± 2 words per category used. However, the slopes in the present study were both significantly less than the minimum of the 5 ± 2 range.

The y -intercept values for Experiments II and III were 23.5 and 19.0 respectively. The values Mandler obtained were 11.2 and 13.9 for his experimentally paced tasks, and the median value for his self-paced tasks was 10.6. He noted that there was a small negative correlation between the slopes of the category-recall functions and the y -intercept values for his seven self-paced tasks. In the present experiments, the category-recall slopes were considerably lower and the y -intercept values considerably higher than those Mandler reported, so the present data conform with that negative correlation.

Mandler (1967) argued that the y -intercept value represents the amount recalled on the basis of some organization other than that represented by the number of categories. If this explanation is accepted, then the subjects in Experiments II and III were not, for the most part, using their sorting categories as the basis of their organization. The question, then, is why Mandler's subjects used the sorting task as a basis for their organization, while the present subjects did not.

Mandler's explanation for the magnitude of the y -intercept and the relationship between slope and y -intercept values is based on the assumption that organization is necessary for recall. The data that make this assumption plausible are the dramatic relationships found between number of categories and recall. However, data from four groups of subjects, two groups used by Nelson et al. (1970) and two used in the present study, have failed to produce the dramatic relationship between number of categories and recall when number of trials and time per trial were experimentally controlled. There appear to be two possible explanations for the discrepancies between Mandler's results and those of Nelson et al. and the present study. One is that organization is necessary for recall but that the sorting task does not ensure that subjects will use

the organization provided by the sorting task as the basis for their recall. The second is that organization is not necessary for recall but that it can improve recall in some situations.

Assigned categories

Although in both Experiments II and III subjects who were assigned to use six categories in the second task recalled slightly more than subjects who were assigned to utilize three categories, in neither case was the difference significant. Only in Experiment III when first-list recall scores were held constant did category assignment even approach significance.

Strictly interpreted, Mandler would predict subjects to recall 15 ± 6 words more when assigned to use six categories than to three categories. In the present study, the differences were on the order of only 1-3 words. Although the data do not necessarily warrant accepting the null hypothesis, the effect, if there is one, is small.

The considerably larger effects that Mandler (1967, 1968) reported when subjects were assigned their number of categories might have been due to his failure to control the number of trials and the time per trial. Alternatively, the explanations that have been advanced to explain the discrepancies in the magnitude of the effect when subjects were allowed to choose their number of categories could apply. That is, either Mandler's subjects were using the organization provided by the sorting task and the subjects in the present study were not using that organization, or Mandler's subjects were organizing in order to recall and the subjects in the present study were not organizing but still recalling.

Individual differences

Little could be expected from this part of the investigation, since the prerequisite—large differences between high and low categorizers—was not observed. Experiment I indicated that there were individual differences in categorization. Experiment II indicated that these differences interacted with category assignment to produce differences in second-list recall, such that subjects tended to recall more when assigned a number of categories similar to the number they habitually used. However, the results of Experiment III did not replicate this interaction, although the same overall pattern of results occurred. In Experiment III, there was some indication that with a 'habitual' choice of number of categories (as measured by first-list choice), high categorizers performed better than low categorizers on the second list. These results suggest

that individual differences in categorization may confound the category-recall function, so that it may be inappropriate to draw causal inferences from studies which allow the subjects to choose their number of categories.

The results on individual differences in recall ability are much clearer. In both Experiments II and III, the best predictor of second-list recall performance, regardless of number of first- or second-list categories used, was first-list recall performance. It could be that subjects differ in their recall ability because of differences in the strategies they employ. And one such strategy could involve the manner in which they sort the words into categories. However, it is clear in the present data that the number of categories into which they sorted those words accounted for only a very small proportion of the variation in their recall performance.

Notes

This research was supported by Grant APA337 from the National Research Council of Canada to the second author. Received for publication May 13, 1971.

References

- Draper, N. R., and Smith, H. 1966. *Applied regression analysis*. New York: Wiley.
- Earhard, M. 1967. Subjective organization and list organization as determinants of free-recall and serial-recall memorization. *Journal of Verbal Learning and Verbal Behavior* 6:501-507.
- Earhard, M., and Endicott, O. 1969. Why are there individual differences in subjective sequential organization during free-recall memorization? *Journal of Verbal Learning and Verbal Behavior* 8:316-319.
- Mandler, G. 1967. Organization and memory. In *The psychology of learning and motivation*, ed. K. W. Spence and J. T. Spence. New York: Academic Press.
- Mandler, G. 1968. Organized recall: Individual functions. *Psychonomic Science* 13:235-236.
- Nelson, A. K., McRae, B. C., and Sturges, P. T. 1970. Recall as a function of instructions and level of organization. Paper presented at the meeting of the Western Psychological Association, Los Angeles, April.
- Paivio, A., Yuille, J. C., and Madigan, S. A. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology* 76(1, part 2).
- Tulving, E. 1962. Subjective organization in free recall of "unrelated" words. *Psychological Review* 69:334-354.

A TEST OF THE RELATIONAL HYPOTHESIS OF PERCEIVED SIZE

Walter C. Gogel and Richard D. Sturm

University of California, Santa Barbara

The relational hypothesis states that the retinal size of a frame directly affects the perceived size of an enclosed line. The context hypothesis states that the perceived size per unit of retinal size of the line and frame will be equal when the line and frame are simultaneously presented at the same apparent distance. The results of two experiments in which the predictions from the two hypotheses differed support the context hypothesis.

The *relational hypothesis of perceived size* states that the perceived size of an object is influenced by the size of surrounding objects. For example, if two lines of the same size are presented in frames of different sizes, then the line presented in the smaller frame will appear larger than the line in the larger frame (Künnapas, 1955). This result is termed a *size-induction effect*. Although this effect is not large when the stimuli are presented in a well-structured visual field, it increases if nothing is visible except the two lines and two frames—presumably because the additional portions of the field provide a more inclusive surround that limits the induction effects of the individual frames (Rock and Ebenholtz, 1959).

A phenomenon similar to size induction occurs in the perception of motion. Brown (1965) was able to summarize the results from a series of experiments on motion perception under the concept of the *transposition* principle. He found that the apparent velocity of a small moving object within a frame depends on the size of the frame and that if all the dimensions of the object and frame are multiplied by a constant (transposed), the physical velocity of the moving object must be multiplied by approximately the same constant in order for the perceived velocity to remain unchanged. Assuming similarity of the processes underlying the perception of size and velocity (Rock, Hill, and Fineman, 1968), the transposition principle can be applied to the perception of both size and velocity, to determine the theoretical values the induction effects would approach if no contrary factors were present. In size induction, for example, if only the induction effect were present, a 1-in. line in a 2-in.

frame should appear equal in length to a 2-in. line in a 4-in. frame.

It has also been assumed that it is the *proximal* size (the retinal size, or visual angle) of the frame that is the important factor in size induction (see Epstein, 1962; Rock and Ebenholtz, 1959). Also, Wallach (1939) suggested that the transposition principle as applied to the perception of motion should be defined at the proximal rather than the distal level. Consistent with the transposition principle defined at the proximal level, the relational hypothesis would state that the perceived size (S'_L) of the enclosed line varies directly with its visual angle (θ_L), and inversely with the visual angle (θ_F) of the enclosing frame. Or

$$S'_L = C\theta_L/\theta_F, \quad [1]$$

where C is a constant of proportionality.

Although Equation 1 refers to a single frame-line combination, it can, of course, be used to predict the ratio of the perceived sizes of lines in two different frames, by applying it to each of the frame-line combinations. To illustrate, suppose that a 3-in. line is presented within a 2-by-4-in. frame and a 9-in. line is presented within a 6-by-12-in. frame, with all of these objects at the same physical distance from the observer. Since the ratio of the visual angles of the two frames is equal to the ratio of the visual angles of the two lines, the prediction from Equation 1 is that the ratio of the perceived sizes of the two lines will be unity—that the two lines will be perceived as equal in size. Note that the predictions from the relational hypothesis can be made using Equation 1 for any visual angles of the lines and frames. For example, if both the lines in the example above were 3 in. long, the prediction from the relational hypothesis would be that the line in the smaller frame will be perceived as three times as large as the line in the larger frame.

Tests of the relational hypothesis have been conducted in situations in which all of the objects (lines and frames, for example) were at the same physical distance from the observer. Under these conditions, the size-induction effects are in the direction predicted by, but not as large as expected from, the relational hypothesis (Künnapas, 1955; Rock and Ebenholtz, 1959). This inability to predict the magnitude of the induction effect should not be interpreted as a limitation of the relational hypothesis, however, since it can instead be explained by the presence of another, often conflicting factor.

This additional factor is the effect of perceived distance on perceived size. The interaction of perceived size and perceived distance is indicated by the *size-distance invariance* hypothesis, which in a simple form can be expressed as

$$S' = K\theta D' . \quad [2]$$

In Equation 2, S' is the perceived size of any object (a line or frame), θ is its visual angle, D' is the perceived distance of the line or frame, and K is a constant of proportionality. If the lines in the two frames are perceptually as well as physically at the same distance from the observer, it follows from Equation 2 that the ratio of the perceived sizes of the lines will be proportional to the ratio of the visual angles of the lines. Since the ratio of the visual angles of lines at a constant physical distance is proportional to the ratio of their physical sizes, it follows that the ratio of the perceived sizes of the lines will be proportional to the ratio of their physical sizes. Thus, according to the hypothesis of size-distance invariance, if the lines are perceived to be at the same distance, then the ratio of their sizes will be correctly perceived with the amount of the size-induction effect zero. In other words, the perception that the lines are at the same distance tends to counteract the size-induction effect expected from the relational hypothesis. Thus, even if the relational hypothesis is valid, two lines that are at the same perceived distance would be expected to show size-induction effects less than those predicted by Equation 1.

On the other hand, the perceived-size effects predicted by Equations 1 and 2 can be in the same direction. This can occur in any situation in which the frame-line configuration with the smaller visual angle appears to be more distant than the frame-line configuration with the larger visual angle. The difference in the perceived distance of similarly shaped objects as a function of visual angle is known as the *relative-size cue to distance*. Consider again the example in which a 3 in. line is presented in a 2-by-4-in. frame and a 9-in. line in a 6-by-12-in. frame, with all the objects at the same physical distance. If the relative-size cue is effective, the retinally smaller configuration (the 3-in. line and 2-by-4-in. frame) will appear at three times the distance of the larger configuration, and by Equation 2, the 3-in. line will be equal in perceived size to the 9-in. line. This is the same result predicted by Equation 1.

It was noted above, though, that the prediction of the perceived size of a line in an enclosing frame will not always be identical from Equations 1 and 2. Assuming that $D'_L = D'_F$, the difference in the predicted value of S'_L from these two equations is $\theta_L (KD'_F - C/\theta_F)$. It follows that the only condition under which Equations 1 and 2 will provide the same predicted value of S'_L (again assuming that the line appears at the distance of its enclosing frame) is when

$$C/\theta_F = KD'_F . \quad [3]$$

That is, Equations 1 and 2 will make the same prediction about the ratio of the perceived sizes of lines in different frames whenever the ratio of the perceived distances of the frames is the inverse of the ratio of the visual angles of the frames. For all other conditions, the predicted ratio will differ from the two equations. Also, if for any reason the smaller frame-line combination is made to appear closer to the observer than the larger frame-line combination, the direction as well as the magnitude of the predictions from the two hypotheses will differ.

The point to be emphasized is that in any test of the relational hypothesis, perceived distance can contribute to perceived size and, except when $C/\theta_F = KD'_F$, can modify the results expected from the relational hypothesis. That is, the effect of perceived distance on perceived size according to Equation 2 is always present and, except in the circumstances defined by Equation 3, is such as to modify the perceived results expected from the relational hypothesis. Thus, a finding that induction effects of the magnitude expected from Equation 1 have not occurred is not acceptable evidence against Equation 1, nor is it unexpected in terms of the known relations that occur between perceived distance and perceived size.

But if size-distance invariance provides a result that often interferes with achieving the magnitude of results expected from Equation 1, the converse is also true. If the relational hypothesis is valid, it will provide a result which, under all conditions except those expressed by Equation 3, will oppose the predictions expected from the hypothesis of size-distance invariance. Thus, a general test of the relational hypothesis would be to determine whether the size perceptions of lines in frame-line combinations do differ from Equation 2 in the direction expected from Equation 1.

Fortunately, from the assumption that the enclosed line appears at the distance of the enclosing frame in each of the frame-line configurations, a more restricted form of Equation 2 can be used for this test of the relational hypothesis. This restricted form of Equation 2 is called the *context hypothesis of perceived size* (Gogel, 1971; Gogel and Mershon, 1968), which states that simultaneously presented objects at the same perceived distance and similar directions have equal values of perceived size (S') per unit of retinal size (θ). Consider the case of a frame (F) and an enclosed line (L). From Equation 2, it follows that

$$D'_F - D'_L = 1/K(S'_F/\theta_F - S'_L/\theta_L), \quad [4]$$

where S'_F/θ_F and S'_L/θ_L are the perceived sizes per unit of retinal size and D'_F and D'_L are the perceived distances of the frame and line re-

spectively. If the line and enclosing frame appear at the same distance from the observer, then $D'_F - D'_L = 0$ and

$$S'_L/\theta_L = S'_F/\theta_F. \quad [5]$$

That is, Equation 5, the context hypothesis, states that the ratio of the perceived sizes of the line and enclosing frame is equal to the ratio of their retinal sizes. But although Equation 5 is derived as a special case of Equation 2, it is not necessarily *limited* to situations in which Equation 2 can be shown to be valid. In order to apply the context hypothesis to the size-induction effect, the only requirements are that the objects (in this study, a frame and enclosed line) be presented simultaneously in similar visual directions at the same apparent distance. And in order to apply Equation 5 to the perceived sizes of two lines in two frame-line combinations of different size, all that is required is that each line appear at the distance of its enclosing frame. It is not required that the two frames or two lines appear at equal distances.

Comparing Equations 1 and 5, the difference in the predicted size of the enclosed line is given by $(\theta_L/\theta_F)(S'_F - C)$. In other words, the relational and context hypotheses will make identical predictions only when

$$S'_F = C. \quad [6]$$

Or, considering different sizes of frame-line configurations, Equations 1 and 5 will make predictions of the same magnitude and direction only when the perceived sizes of the differently sized frames are the same.

Thus, the purpose of the present study was to test the relational hypothesis (Equation 1) by testing the ability of Equation 5 to predict the perceived sizes of lines enclosed in frames of different retinal sizes. If Equation 1 also applies under these circumstances, it will provide a source of the perceived sizes of the enclosed lines which, except when $S'_F = C$, differs from that expected from Equation 5. The observer will thus be faced with differing information about S'_L from these two sources, and the resulting compromise of these differing sources should produce a modification of his responses from those predicted by Equation 5. If, on the other hand, Equation 5 is completely valid under conditions in which a size-induction effect is expected to occur, the contribution of the retinal size of the frame to the induction effect independent of the perceived size of the frame (Equation 1) would be expected to be negligible.

EXPERIMENT I

Rock and Ebenholtz (1959) were interested in eliminating the effect of the relative-size cue to distance between the physically equidistant

frame-line configurations of different retinal size. For this reason, in most of their experiments on the size-induction effect the viewing was binocular in order to introduce the distance cue of convergence, and the different configurations were separated directionally by 180 deg. Although binocular as compared with monocular observation is useful for this purpose (Gogel and Sturm, 1972), a directional separation of 180 deg as compared with 0 deg does not seem to reduce the effectiveness of the relative-size cue (Gogel and Sturm, 1971).

It was not considered to be necessary in the present experiments to eliminate the relative-size cue between the successive presentations of the frame-line configurations, because the relational and context hypotheses do not predict the same results except when the frames of different visual angle appear to have the same size (Equation 6). It follows that either monocular or binocular observation provides a test of the relational hypothesis by testing the validity of the context hypothesis if it is found that the same perceived size did not occur with the different frames. In general, though, where consistent with the purposes of this study, the stimulus conditions used were similar to those used by Rock and Ebenholtz (1959).

Method

—Apparatus—An observation booth was constructed, in which the observer sat midway between two observation positions separated by 180 deg (see Figure 1, in Gogel and Sturm, 1971). Each observation position was totally dark and contained an adjustable head-and-chin rest. The stimuli, consisting of frames and enclosed lines, were monocularly viewed through a nonrestrictive aperture, with the remainder of the visual field totally dark. Two sizes of luminous rectangular frames— $5\frac{3}{4}$ by $11\frac{3}{4}$ in. and $1\frac{1}{2}$ by $3\frac{1}{2}$ in., with the luminous sides of the frames $\frac{1}{4}$ in. wide—were constructed by masking electroluminescent panels and were presented with the longest side vertical. These luminous frames approximated the dimensions of the frames used by Rock and Ebenholtz (1959), namely, 6 by 12 in. and 2 by 4 in. As in that study, the centers of the frames were on the observer's line of sight at a distance of 5 ft. The luminous vertical line enclosed within each frame originated at the center of the base of the frame. The line within the larger frame was $6\frac{1}{16}$ in. high and $\frac{1}{16}$ in. wide, and the line within the smaller frame was $2\frac{1}{16}$ in. high and $\frac{1}{4}$ in. wide. The luminance of all the stimuli was .04 ftL. The ratio of the large to the small line is approximately the average ratio (2.2:1) of the large line in the large frame required to match a 3-in. standard line in the small frame of Rock and Ebenholtz's study.

—Observers—The observers were 80 men and women enrolled in an introductory course in psychology. All were naive as to the purpose of the experiment, and all had a near and far visual acuity of 20/20, corrected if necessary.

—Procedure—Each observer was led blindfolded into the dark observation booth. The blindfold was removed and, after a 5-min period of dark adaptation, the observer was instructed to place his head in one of the observation positions and was shown the first stimulus. He was then asked to indicate the following, with all height or distance responses expressed in feet or inches or in some combination of feet and inches: (a) the height of the enclosed vertical line, from the top of the line to where it ended at the bottom of the rectangle, (b) the height of the rectangle, (c) whether the line and the rectangle were at the same or different distances from his eyes and, if different, which appeared closer and by what amount, and (d) the distance of the rectangle from his eyes. The shutter was then closed, and the observer turned 180 deg to the other observation position. After 15 sec, he was instructed to place his head in the second observation position, the second stimulus was presented, and he again made distance and height responses. Half the observers were presented with the small frame and line first and the large frame and line second. For the 40 other observers, the order in which the small frame-line configuration and the large frame-line configuration were presented were reversed. The two presentations to an observer were always separated by 180 deg, with the observation position used for the first presentation counterbalanced across observers.

Results

The results of Experiment I are shown in Table 1. The perceived distance, D' , of the line was obtained by adding to the reported distance of the enclosing frame the perceived depth (plus or minus) between the line and enclosing frame. Since the distributions were often skewed, the medians, rather than means, may be considered to best represent the data. As shown in Table 1, the mean and median perceived distances to a line and to its enclosing frame were very similar. Of the 80 observers, 61 reported no depth between a line and its enclosing frame on the first presentations; only 5 observers reported a depth greater than 6 in. Of the 80 observers, 68 reported no depth between a line and its enclosing frame on the second presentations; only 2 observers reported a depth greater than 6 in. According to the context hypothesis, therefore, the median perceived size per unit of visual angle (S'/θ) of a line and the frame enclosing the line should be very similar.

It will be noted in Table 1 that on first presentations median perceived distances of the differently sized frames were not inversely related to the visual angles of the frames. Using the Mann-Whitney U test, the bulk of these perceived distance scores for the frames were not significantly different when the small rather than the large frame was used ($Z = .3, p > .05$). On the second presentations, however, the bulk of the perceived distances for the small frame were significantly greater than for the large frame ($Z = 3.28, p < .01$). By definition, if the rela-

Table 1. Means, medians, and standard deviations in inches of perceived distance and perceived height obtained from the large and small frame and line, Experiment I

	Line				Frame			
	First presentation		Second presentation		First presentation		Second presentation	
	Large	Small	Large	Small	Large	Small	Large	Small
Perceived distance (D')								
Mean	47.2	65.0	36.3	94.0	47.9	64.4	38.4	92.5
Median	33.0	24.0	13.5	36.1	30.0	24.0	13.5	36.0
Standard deviation	50.9	172.0	85.4	111.9	53.6	172.1	85.7	108.7
Perceived height (S')								
Mean	8.2	3.7	8.7	11.8	13.2	6.8	14.5	14.4
Median	2.0	1.5	2.5	1.5	3.2	2.0	4.0	2.0
Standard deviation	13.7	9.7	16.6	38.4	22.8	17.3	26.8	47.7

tive-size cue occurred in the present experiment, it must have done so between successive presentations. Thus, in order to demonstrate that the relative-size cue to distance was effective between the differently sized frames, the effect of any tendency to perceive the frames at different distances in the first presentations must be removed.

An analysis that adjusts for any such difference in first presentations consists in comparing the obtained difference in D' between second presentations with the obtained difference in D' between first presentations (Gogel, 1969; Gogel and Sturm, 1971). It was found, using the Mann-Whitney U test, that the bulk of the D' differences $A_2 - B_2$ were greater than those $A_1 - B_1$ at the .01 level ($Z = 2.43$), where A and B refer to the small and large frames and 1 and 2 refer to the first and second presentations. Thus, in agreement with the previous study (Gogel and Sturm, 1972), the results support the conclusion that the relative-size cue can occur between presentations separated by 180 deg. As a consequence of the relative-size cue to distance, the retinally smaller frame (and line) appeared to be more distant than the retinally larger frame (and line) in the successive presentations.

But despite the relative-size cue to distance, the small frame was perceived as smaller than the large frame in the second presentations. The tendency to see the small frame as smaller than the large frame on both first and second presentations was significant at the .01 level

($Z = 3.71$ and 2.88 respectively). It follows that Equation 6 was not satisfied and that Equations 1 and 5 therefore predict different results in this experiment.

The degree to which the obtained median results satisfy the context hypothesis can be seen from Figure 1, in which the S'/θ of a line is plotted against the S'/θ of its enclosing frame, where θ is the visual angle of the height of the line or frame expressed in radians and the S' is the perceived height of the line or frame in centimeters. The filled circles are from Experiment I, with A and B referring to the small and large frames (or lines) and 1 and 2 referring to first and second presentations. The dashed line is the prediction from the context hypothesis, with which the data are obviously in close agreement. This close fit to the context hypothesis applies to individual as well as median data, in that the Pearson product-moment correlations of the S'/θ of the line and enclosing frame were .99 or 1.00 for both the small and large line-frame configurations on both the first and second presentations.

Some caution is necessary in interpreting these unusually high correlations. They might partly be due to a multiplicative constant involved in attaching numbers to perceived sizes, with this constant differing between observers. For example, one observer might tend to give consistently larger reports of size on both frame and line than another ob-

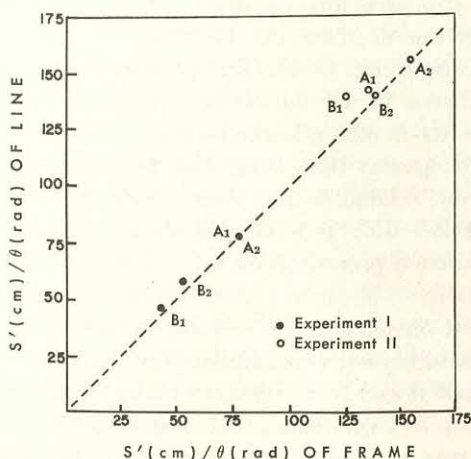


Fig. 1. Relation between the perceived size (S') per unit of visual angle (θ) of the lines and the enclosing frames. The dashed line is the relation predicted by the context hypothesis of perceived size. A and B refer to the small and large frames (or lines); 1 and 2 to the first and second presentations.

server, even though the perceived sizes were the same for the two observers. This difference would contribute to the value of the correlation coefficients, regardless of the validity of the context hypothesis.

To examine this last possibility, it was noted that the context hypothesis would predict the ratio of the S'/θ values of the frames from the first and second presentations to be positively correlated with (equal to) the ratio of the S'/θ values of the enclosed lines. Since ratios of S'/θ are being used, the resulting r cannot be attributed to observer differences in scaling reports of size. These ratios of S'/θ obtained for each observer resulted in an r of $+.99$ ($df = 78, p < .01$).

EXPERIMENT II

Experiment II was identical to Experiment I in all respects except that binocular rather than monocular observation was used, to provide a range of conditions against which the relational and context hypotheses could be compared. Eighty new observers were used.

Results

The results from Experiment II are shown in Table 2. As in Experiment I, the lines appeared approximately at the distance of their enclosing frames. Of the 80 observers, 68 reported no depth between a line and its enclosing frame on the first presentations; only 2 reported a depth greater than 6 in. Of the 80 observers, 69 reported no depth between a line and its enclosing frame on the second presentations; only 4 reported a depth greater than 6 in. The perceived distances, D' , of the frames were not related to the visual angle of the frames on the first presentations ($Z = 1.69, p > .05$) but were inversely related to visual angle on the second presentations ($Z = 2.36, p < .01$). The bulk of the D' differences $A_2 - B_2$ were greater than the bulk of the D' differences $A_1 - B_1$ at the .05 level ($Z = 1.94$). It follows that the relative-size cue was effective between successive presentations.

But, it will be noted, the D' differences between successive presentations were much less in Experiment II than in Experiment I. Only when the large frame was presented first (not when the large frame was presented second) were the bulk of the perceived distance scores greater for the small than the large frame on the successive presentations ($Z = 2.64, p < .01$, and $Z = 1.36, p > .05$, respectively). As expected (Gogel and Sturm, 1972), the relative-size cue between successive presentations

Table 2. Means, medians, and standard deviations in inches of perceived distance and perceived height obtained from the large and small frame and line, Experiment II

	Line				Frame			
	First presentation		Second presentation		First presentation		Second presentation	
	Large	Small	Large	Small	Large	Small	Large	Small
Perceived distance (D')								
Mean	57.4	44.6	52.8	80.4	58.4	44.2	53.0	80.3
Median	40.0	36.0	30.0	48.0	43.5	36.0	36.0	48.0
Standard deviation	62.7	38.7	64.0	131.0	61.8	38.6	59.0	131.0
Perceived height (S')								
Mean	9.0	4.1	10.3	5.2	14.8	5.1	18.0	6.4
Median	6.0	2.8	6.0	3.0	9.5	3.5	10.5	4.0
Standard deviation	9.5	4.5	12.6	8.7	16.1	5.8	21.7	9.7

was more effective with monocular observation (Experiment I) than with binocular observation (Experiment II).

As in Experiment I, however, the median perceived height, S' , of the frame was a direct function of the visual angle of the frame in both the first and second presentations ($Z = 5.1, p < .01$, and $Z = 4.91, p < .01$, respectively). Thus, Equation 6 was again not satisfied, and the predictions from Equations 1 and 5 differ in this experiment also.

The relation between the S'/θ of the lines and the S'/θ of their enclosing frames is shown by the open circles in Figure 1, as computed using the median S' data of Table 2. Except for one data point (B_1) the results are again very close to the predictions from the context hypothesis of perceived size. The four Pearson product-moment correlation coefficients between the S'/θ value of the lines and enclosing frames for the large and small frames on the first and second presentations were all .98 or higher. Also, the correlation between the ratios of S'/θ values for the lines and ratios of S'/θ values for the frames from the first and second presentations was .98.

DISCUSSION

Since the perceived sizes of the differently sized frames were not the same in either experiment, Equation 6 was not satisfied and the pre-

dictions from the relational and the context hypotheses differ. Therefore, different values of information about the perceived sizes of the enclosed lines were available from these two hypotheses, assuming that both are valid.

The relational hypothesis was not expected to predict completely the perceived size of the lines, of course, but by the reasoning outlined at the beginning of this paper, it should, if valid, be capable of reducing the predictive ability of Equation 5 for objects enclosed in frames. That the predictions from the relational hypothesis have no discernible effect on the data is shown by Figure 1, in which the S'/θ from the frames and enclosed lines is not only related but very close to the predicted 45-deg line. Only if the line through the data of Figure 1 did not have a slope of unity or were curvilinear, and only if the linear correlations between S'/θ from the frame and line were appreciably different from unity, could it be concluded that some factor other than the context hypothesis must be postulated in order to explain the results.

It will be noted also that, except possibly in the case of B_1 in Experiment II, the context hypothesis applied to the results of the first as well as the second presentations; that is, it applied equally well whether or not the relative-size cue between the frames was present to contribute to the perceived size of the enclosed line.

From Equation 1, the relational hypothesis predicts that the S'/θ of the enclosed line will be inversely proportional to the visual angle of the enclosing frame. The relevant data here are presented in Figure 2, which plots the S'/θ of the line against $1/\theta$ of the enclosing frame. The larger values of S'/θ in Experiment II as compared with those in Experiment I (see Figures 1 and 2), in agreement with the hypothesis of size-distance invariance (Equation 2), were at least partially the result of the greater (more veridical) perceived distance from binocular observation. It should be noted, however, that in each experiment there was a consistent tendency for the S'/θ of the line to *increase* with an increase in $1/\theta$ of the enclosing frame.

From Equation 6 the predictions from the context and relational hypotheses are the same when $C = S'_F$, that is, when $C/\theta_F = S'_F/\theta_F$. Since the data from both experiments as shown in Figure 1 support the context hypothesis, the tendency for S'_L/θ_L in Figure 2 to increase with $1/\theta_F$ must have been the result of an equal tendency for S'_F/θ_F to increase with $1/\theta_F$. Partially, this can be accounted for by the changes in perceived distance that occurred between successive presentations. Thus, the increase in S'_L/θ_L with increasing $1/\theta_F$ in Figure 2 between

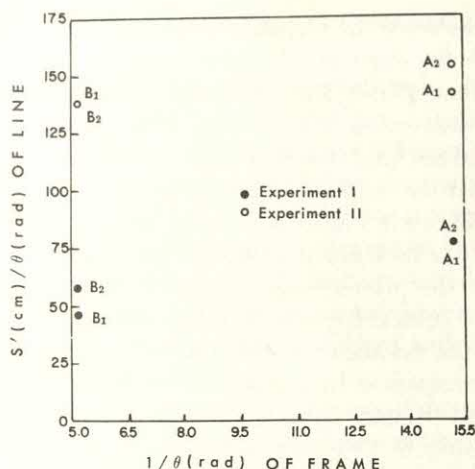


Fig. 2. Relation between the perceived size (S') per unit of visual angle (θ) of the enclosed lines and the reciprocal of the visual angle ($1/\theta$) of the enclosing frames. If the obtained data were completely consistent with the relational hypothesis of perceived size, all the points would be on a straight line. A and B refer to the small and large frames (or lines); 1 and 2 to the first and second presentations.

B_1 and A_2 or between B_2 and A_1 would be expected from the relative-size cue to distance.

On the other hand, the relation between S'_L/θ_L and $1/\theta_F$ occurring with A_1 and B_1 in Figure 2 remains to be explained for the two experiments. The small line and frame appeared somewhat nearer (but not significantly nearer) than the large line and frame on the first presentations in both experiments. Nevertheless, the ratio of the perceived size of the line in the small frame to the perceived size of the line in the large frame was greater than the ratio of the physical sizes (or visual angles) of the two lines. This is the type of phenomenon that is usually considered as illustrating the size-induction effect and is consistent with Rock and Ebenholtz's (1959) conclusion that the effect can occur even though the smaller frame-line configuration does not appear at a greater distance than the larger frame-line configuration.

This last result is inconsistent with the hypothesis of size-distance invariance (Equation 2), which would require that the perceived distance of the line with the greater S'/θ value (the small line) be greater than the perceived distance of the line with the lesser S'/θ value (the large line). As Figure 1 shows, however, the results from B_1 and A_1 even in

these circumstances were, in general, consistent with the context hypothesis. Although the slope of a line connecting A_1 and B_1 in Experiment II was less than 45 deg, the slope of a line connecting A_1 and B_1 in Experiment I was very close to 45 deg. Also, it must be emphasized that Equation 6 did not hold in either experiment between the different frame sizes and that the relational hypothesis should have resulted in a slope of less than 45 deg in Figure 2 between the large and small frame-line configurations for both first and second presentations.

In other words, the present study supports the conclusion that, in agreement with the context hypothesis, wherever a so-called size-induction effect occurs, the modification of the perceived size of the enclosed object (line) is accompanied by a modification in the perceived size of the enclosing object (frame). Since the enclosing frame is not itself enclosed by another object, this latter effect can hardly be considered a size induction.

The context hypothesis predicts that the S'/θ value of the frame and enclosed line that appear at the same distance will be the same. It does not predict that the S'/θ of the two frame-line configurations will necessarily be equal even though they appear at the same distance. The reason for this is that the context hypothesis requires that the frame and line to which it applies be presented simultaneously and in similar visual directions, as well as that the frame and line be at the same perceived distance from the observer. The results obtained on the first presentations of the different retinal sizes (A_1 and B_1) in the present experiments were not in agreement with the hypothesis of size-distance invariance, even though they were consistent with the context hypothesis of perceived size. Since the hypothesis of size-distance invariance contains no requirement that the objects to which it applies be simultaneously presented in similar directions, the context hypothesis, although related to Equation 1, is not totally a special case of Equation 1. Instead, the context hypothesis may be an example of the adjacency principle (Gogel, 1971).

The *adjacency principle* states that the perceptual interactions between objects decrease with an increase in their perceived separation in depth, direction, or time. Thus, the greater the relative adjacency of simultaneously presented objects in perceived space (perceived depth and direction), the more the S'/θ values of the objects will be similar. This is equivalent to stating that the S'/θ of a local portion of perceived space tends to be homogeneous. Thus, if a modification in S'/θ occurs with respect to the enclosing frame, it must also be present with respect to the enclosed line.

Consider now the case in which two frame-line configurations are presented simultaneously at the same distance but in *different* visual directions. If binocular observation is used, it is likely that the two configurations will appear at the same distance. Assume also that under these conditions the two configurations have different values of S'/θ despite appearing at the same distance. When the perceived sizes of the lines are considered, this is an example of the size-induction effect. In this case, however, the S'/θ values in the vicinity of the two frames differ. If the line enclosed by one of the frames is then moved in the frontoparallel plane away from that frame and toward the other frame, it will move through a gradient of S'/θ values and should change appropriately in its perceived size.

This last example is an illustration of a two-dimensional illusion and suggests that many two-dimensional illusions can be specified as a gradient of S'/θ across the visual field. If the gradient of S'/θ differs in the horizontal and vertical directions, a line oriented so as to have both a horizontal and vertical component will have an illusory slant. In all these cases, the context hypothesis should predict the perceived length or orientation of a line moved about in the frontoparallel plane as a function of the distribution of S'/θ values in that plane.

It might be noted here that the context hypothesis describes the perceptual consequences for objects (such as lines) of being in different portions of the visual field having different values of S'/θ . It does not describe the factors that produce S'/θ differences between different parts of the visual field. Probably there are a variety of such determining factors, with the perceptual consequences of the differences in S'/θ independent of the particular factor that produced those differences.

Some interesting questions result from applying the context hypothesis to two-dimensional illusions. Suppose that a difference in S'/θ is produced in the apparent frontoparallel plane, for example, by binocularly observing horizontally separated but physically equidistant portions of the Müller-Lyer illusion. What is the gradient of S'/θ between the shafts of the two portions of the illusion? Will the S'/θ of the probe line equally displaced from the two shafts be the average of the S'/θ values at the shafts?

Questions of this type are probably important in predicting the distortion in perceived shape of an extended visual stimulus (such as a large square) situated in a frontoparallel plane containing a nonuniform distribution of values of S'/θ . Furthermore, it seems clear that the problem of what determines a distribution of S'/θ in an apparent frontoparallel plane is but a special case of the more general problem of the

factors that determine a distribution of S'/θ in apparent three-dimensional space. Thus, in addition to making the relational hypothesis (and perhaps the transposition principle) unnecessary, the context hypothesis may well be useful in the descriptive analysis of a variety of situations involving visual illusions.

Notes

This investigation was supported by PHS Research Grant MH15651 from the National Institute of Mental Health. The authors wish to thank Richard Weinthal for his help in the collection of the data. Received for publication July 14, 1971.

References

- Brown, J. F. 1965. The visual perception of velocity. In *Readings in the study of visually perceived movement*, ed. I. M. Spegel, pp. 64-107. New York: Harper & Row.
- Epstein, W. A. 1962. A test of two interpretations of the apparent size effects in a distorted room. *Journal of Experimental Psychology* 63:124-128.
- Gogel, W. C. 1969. The absolute and relative size cue to distance. *American Journal of Psychology* 82:229-234.
- Gogel, W. C. 1971. The adjacency principle and three-dimensional visual illusion. In *Psychonomic Monographs*, Human space perception: Proceedings of the Dartmouth conference, ed. J. Baird.
- Gogel, W. C., and Mershon, D. H. 1968. The perception of size in a distorted room. *Perception and Psychophysics* 5:7-9.
- Gogel, W. C., and Sturm, R. D. 1971. Directional separation and the size cue to distance. *Psychologische Forschung* 35:57-80.
- Gogel, W. C., and Sturm, R. D. 1972. A comparison of accommodative and fusional convergence as cues to distance. *Perception and Psychophysics* 11:166-168.
- Künnapas, T. M. 1955. Influence of frame size on apparent length of a line. *Journal of Experimental Psychology* 50:168-170.
- Rock, I., and Ebenholtz, S. 1959. The relational determination of perceived size. *Psychological Review* 66:387-401.
- Rock, I., Hill, A. L., and Fineman, M. 1968. Speed constancy as a function of size constancy. *Perception and Psychophysics* 4:37-40.
- Wallach, H. 1939. On constancy of visual speed. *Psychological Review* 46:541-552.

TRANSFER IN THE A-B', A-B PARADIGM AS A FUNCTION OF ANTICIPATION-INTERVAL LENGTH

Michael Maus

University of Iowa

The length of the anticipation interval during list 2 learning was 1.25 or 3.00 sec for groups learning paired associates under the A-B', A-B or C-D, A-B paradigms; list length was 15 pairs. The results support a mediational interpretation in that positive transfer for AB' appeared to be a direct function of anticipation-interval length. The concept of the AB' paradigm as a functionally mixed AB' and AC paradigm is discussed.

The A-B', A-B paradigm (henceforth referred to in this paper as AB') consists of two successively presented lists of paired associates; the corresponding stimulus terms in each list are identical and the corresponding response terms are somehow related, usually along some dimension of meaningful similarity. Several hypotheses have been advanced to explain the positive transfer obtained for the AB' paradigm in relation to a C-D, A-B control (henceforth referred to as CD). One is the mediation hypothesis (Barnes and Underwood, 1959), which proposes that, having learned the association A-B' during list 1 learning, the subject uses this association, plus the presumed prelaboratory-learned associative connection between the B' and B responses, to form the associative chain A-B'-B as an aid in his learning the A-B pairs presented in list 2.

One method to determine whether mediation is a reasonable explanation of the positive transfer in the AB' paradigm has been to examine the effects of anticipation-interval length on performance in the AB' paradigm during list 2 learning. It has been proposed that during list 2 learning the utilization of the A-B'-B associative chain requires the elicitation of the implicit associative response B' prior to the actual response B (Richardson, 1967). The occurrence of this intervening process between the presentation of the stimulus and the elicitation of the overt response should increase the time required for the overt response to be made. If mediation does occur in the AB' paradigm, then manipulating the length of the anticipation interval during list 2 learning—and consequently, the time available to mediate—should differentially affect performance in the AB' paradigm and in a nonmediational paradigm such as CD. The

shorter the anticipation interval, the less should be the amount of positive transfer for AB' in relation to CD.

Of those studies involving such a manipulation in the AB' and CD paradigms, two have found a significant interaction between paradigm and length of anticipation interval (Richardson, 1967b; Richardson, 1968), and three have failed to demonstrate any interaction between paradigm and anticipation interval (Runquist and Marshall, 1963; Spear, Mikulka, and Podd, 1966; Richardson, 1967a). Of those studies attempting to relate the B' and B items along some dimension of meaningful similarity such as synonymity, only Richardson (1968) has found a significant interaction between the amount of positive transfer for AB' and the length of the anticipation interval.

The present study was a further attempt to demonstrate that the amount of positive transfer found for the AB' paradigm in which the B' and B terms are meaningfully similar is, in fact, a function of anticipation-interval length. It was noted that in many of the previous studies involving the AB' paradigm and a manipulation of the anticipation-interval length, the positive transfer expected for AB' relative to CD was either small or failed to obtain. To increase the likelihood of obtaining substantial amounts of positive transfer for AB' in the present experiment, a larger number of items per list was used than has usually been employed in similar experiments—this because pilot work showed that by increasing the number of items in the lists, the initial positive transfer for AB' relative to CD could be amplified considerably.

METHOD

—Design—Subjects learned two paired-associate lists representing either the AB' or CD paradigm. Anticipation intervals of either 1.25 or 3.00 sec were used for list 2 learning. The four possible combinations of the two paradigms (AB' and CD) with the two anticipation intervals (1.25 and 3.00 sec) formed the four experimental conditions, labeled AB'-1.25, AB'-3.00, CD-1.25, and CD-3.00.

—Subjects—A total of 80 paid college-age students participated in the experiment. Of the 80, 16 were dropped—5 due to mechanical failures of the experimental apparatus, and 11 due to their failure to meet criterion of $\frac{13}{15}$ correct on list 1 within 20 trials. The numbers who failed to meet criterion on list 1 in conditions AB'-1.25, AB'-3.00, CD-1.25, and CD-3.00 were 5, 1, 0, and 5 respectively. A discussion of the possible differential effects of the unequal numbers of subjects dropped in each of the four conditions is in the section on results.

The remaining 64 subjects were assigned randomly to each of the eight conditions (two different lists were used for each of four experimental conditions) with the restriction that the N th subject not be assigned to a condition until $N - 1$ subjects had been assigned to all other conditions. In addition,

within each of the eight conditions half the subjects were male and half were female.

—Materials—Four 15-item paired-associate lists, designated A₁-B₁, A₁-B'₁, A₂-B₂, and A₂-B'₂, were constructed. The stimuli were trigrams selected from Glaze (1928) at the 47%–53% association level. The response terms were two-syllable adjectives selected from Haagen (1949). The frequency ratings of the adjectives were comparable for all four lists ($.54 \leq X \leq .58$). Intralist similarity was minimized as much as possible in all lists. However, due to the limited number of items with the desired high level of rated similarity within Haagen and the large number of items needed to construct the two sets of lists (30 adjective pairs altogether), some intralist similarity was unavoidable among the response terms, and such items as *horrid*, *gloomy*, *dirty*, and *rotten* appeared within the same lists.

The lists were constructed in such a way that the A₁-B'₁' and A₁-B₁ lists had the same trigrams for stimuli and that their corresponding responses were similar in meaning (the B'₁' and B₁ adjectives had a mean *M* value of 1.27; Haagen, 1949). The A₂-B'₂' and A₂-B₂ lists had similar relationships (the B'₂' and B₂ adjectives had a mean *M* value of 1.29). There was no apparent relationship between either the A₁-B'₁' list and the A₂-B₂ list, or the A₂-B'₂' list and the A₁-B₁ list. The combinations of the various lists were partially counter-balanced such that half the subjects in the CD conditions received A₁-B'₁' as list 1 and A₂-B₂ as list 2, and the other half received A₂-B'₂' as list 1 and A₁-B₁ as list 2. Correspondingly, half the subjects in the AB' condition received A₁-B'₁' as list 1 and A₁-B₁ as list 2, and the other half received A₂-B'₂' as list 1 and A₂-B₂ as list 2.

—Procedure—Subjects were required to learn two successive paired-associate lists using the standard anticipation-learning method. Four different orders of the items were used for each of the lists to prevent serial learning. A Stowe memory drum connected to two Hunter interval timers was used to present the lists. The subjects were instructed not to respond during the first anticipation trial. All learned list 1 at a 2:2-sec rate to a criterion of $1\frac{3}{15}$. A 4-sec intertrial interval was used. If a subject failed to attain a criterion of $1\frac{3}{15}$ after 20 trials (discounting the first trial, on which subjects were instructed not to respond), he was dropped from the experiment. This procedure was used to facilitate the scheduling of subjects.

Before list 2 learning, the subjects were read brief instructions telling them of the particular change in the length of anticipation intervals (1.25 or 3.00 sec) appropriate to the condition they were in. As in list 1 learning, a 2-sec study interval was used. Four different starting orders were employed. Performance was measured over ten trials of list 2 learning. Again a 4-sec intertrial interval was used in all conditions.

RESULTS

List 1 learning

The means and standard deviations of the number of trials to criterion for list 1 learning (ignoring the first trial, on which the subjects were in-

structed not to respond) are shown in Table 1. An analysis of variance revealed no significant effects for list 1 learning [$.14 \leq F \leq 1.77$ (1, 56), $p > .05$ in all cases]. Thus, no significant differences were evident between the groups before their list 2 learning.

Table 1. Means and standard deviations for list 1 and list 2 learning

Condition	List 1 Trials to criterion		List 2 Mean number correct (5 trials)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
AB'-1.25	10.06	4.27	32.56	10.34
AB'-3.00	10.38	4.04	57.25	9.48
CD-1.25	12.75	4.68	27.38	13.98
CD-3.00	10.13	4.03	40.00	12.30

As we mentioned, there was a possibility that differential subject selection occurred, since different numbers of subjects were dropped from the various conditions for failing to meet the criterion of $1\frac{3}{5}$ within 20 trials during list 1 learning ($0 \leq N \leq 5$ per group). However, since there were no significant effects in the analysis of list 1 learning, and since the lists were counterbalanced, there is no reason to believe that the differential loss caused the average ability level of the subjects to be unequal among the four conditions.

List 2 learning

Performance was measured over ten trials of list 2 learning. Although substantial differences in performance as a function of anticipation-interval length continued to be present over all ten trials of list 2 learning, any differences in performance as a function of paradigm disappeared after the first five trials. Because transfer effects are usually most pronounced in the early stages of list 2 learning and have a tendency to dissipate during later trials, performance was analyzed for only the first five trials (again discounting the initial trial, on which the subjects were instructed not to respond).

That performance in the four experimental conditions over the first five trials of list 2 learning is presented in Figure 1. And the means and standard deviations for the total number of correct responses during the first five trials of list 2 learning in each condition are shown in Table 1. The crucial interaction, of paradigm and anticipation-interval length, was found to be significant [$F = 4.14$ (1, 56), $p < .05$].

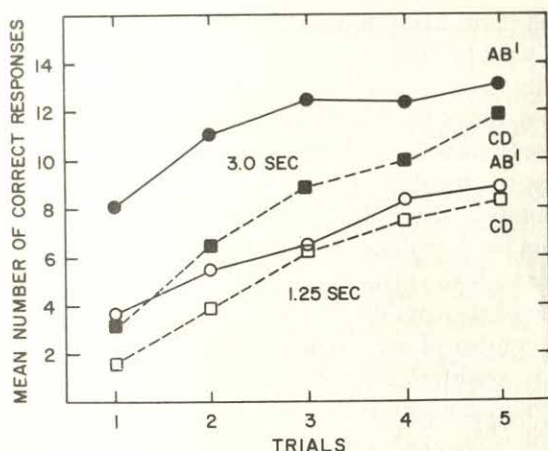


Fig. 1. The mean number of items correct in the four experimental conditions for the first five trials of list 2 learning

For the present experiment, then, the amount of positive transfer for AB' relative to CD was shown to vary as a function of the length of the anticipation interval. As expected, both main effects, of paradigm [$F = 14.77 (1, 56)$] and of anticipation length [$F = 41.40 (1, 56)$], were highly significant ($p < .001$). Although lists did prove to be a significant source of variance [$F = 5.08 (1, 56)$, $p < .05$], none of the interactions between lists and the other main effects proved significant ($F < 1$ in all cases).

The number of occurrences of various types of errors on list 2 in each of the four conditions is presented in Table 2. Immediately to the right of each listing of the actual number of errors is an entry expressing the number of errors as a percentage of the total number of responses for that particular condition. The first column includes those errors where the subject gave the appropriate list 1 response to the stimulus term in place of the correct list 2 response. This type of error could only occur

Table 2. Errors during list 2 learning

Conditions	Appropriately paired intrusions		Intrusions		Errors		Total responses
AB'-1.25	39	5.91%	52	7.89%	138	20.94%	659
AB'-3.00	15	1.49%	21	2.09%	91	9.09%	1,001
CD-1.25			3	.56%	91	17.20%	529
CD-3.00			6	.77%	138	17.73%	778

in the AB' conditions, since only in the AB' conditions were the same stimulus items used in both lists 1 and 2. The second column lists the number of intrusions, which includes all those instances that a list 1 response was given during list 2 learning. For the AB' conditions this number includes those instances presented in the first column. The third column contains the number of overt errors: intrusions from list 1, responses not found in either list 1 or 2, and incorrect responses from elsewhere within list 2. It does not include omissions. The total number of overt responses, both correct and incorrect, is presented in the last column for each of the conditions.

The relative number of occurrences of the various types of errors in the conditions is consistent with the mediation hypothesis. There were substantially more intrusions from list 1 in the AB' condition (73) than in the CD condition (9), with a majority of the AB' intrusions being appropriately paired intrusions (54 out of 73). More importantly, appropriately paired intrusions accounted for 5.91% of the total responses in the AB'-1.25 condition and for only 1.49% of the total responses in the AB'-3.00 condition, which suggests that as the time to mediate is shortened, the mediational chain is interrupted, causing an increase in the likelihood that the mediator rather than the mediated response will be elicited.

DISCUSSION

By demonstrating positive transfer for the AB' paradigm to be a function of anticipation-interval length, the results of the present experiment support the mediational interpretation of the AB' paradigm. They stand in agreement with the results of two studies (Richardson, 1967b; Richardson, 1968) and in disagreement with three others (Runquist and Marshall, 1963; Spear et al., 1966; Richardson, 1967a).

The question that must be asked is why differing results were obtained in the several studies on the effects of anticipation-interval length on AB' transfer. A difference in types of materials does not seem to account for the discrepancy, since only the study by Richardson (1967b) used substantially different materials (letters were used for responses).

Richardson (1968) attributed the inconsistent findings to the difference in lengths of study intervals employed in the various studies. The two studies using a 1-sec study interval (Richardson, 1967b; Richardson, 1968) found that positive transfer for AB' was a function of anticipation-interval length, and those studies using a 2-sec study interval failed to find such a relation between AB' transfer and length of anticipation

interval. Richardson proposed that a 'discovery and utilization' concept of performance in AB'—a concept similar to the one proposed by Schulz and Lovelace (1964) to account for effects of test-interval length in three-stage mediation paradigms—might suffice to explain the conflicting results.

However, as Postman has pointed out (1972), the application of Richardson's 'discovery and utilization' hypothesis requires the occurrence of a triple interaction between length of study interval, length of anticipation interval, and the amount of positive transfer for AB'. A study involving a factorial manipulation of study and test intervals using 1.0- and 2.5-sec study intervals and 1.5- and 3.0-sec test intervals failed to obtain such an interaction for the A-B, B-C, A-C mediation paradigm (Schulz and Weaver, 1968). In addition, the present experiment found the amount of positive transfer for AB' to be a function of anticipation interval length using a 2-sec study interval—the same study interval used in those previous studies that failed to find such an effect. Thus, the difference in lengths of study intervals, although possibly a contributing factor, does not appear to be a sufficient explanation of the conflicting results of the several studies.

A promising clue may lie in the fact that all those studies showing AB' transfer to be a function of anticipation-interval length (Richardson, 1967b; Richardson, 1968; the present experiment) have also shown strong initial positive transfer for AB' in conditions involving long anticipation-interval lengths. In contrast, all those studies finding no such effect of anticipation-interval length on AB' positive transfer (Runquist and Marshall, 1963; Spear et al., 1966; Richardson, 1967a) have failed to obtain substantial amounts of initial AB' positive transfer for *any* length of anticipation interval. It would seem that the success or failure of an experiment in demonstrating mediational effects for the AB' paradigm depends on whether or not strong initial positive transfer is obtained for AB' in those conditions involving long anticipation intervals.

One way to account for the observations above is by a 'mixed list' conceptualization of the AB' paradigm. If the A-B association is lost or if the assumed B'-B relationship is not apparent to the subject, then the nominal AB' paradigm functionally becomes the negative transfer paradigm A-C, A-B (henceforth referred to as AC). If this happens for a certain proportion of the items in a list of nominal AB' items, the list becomes functionally a 'mixed' AB' and AC list. If the effect of anticipation-interval length on AC performance relative to CD is different than for AB', either in degree or in kind, the outcome of any experiment on

the effects of anticipation-interval length on the nominal AB' paradigm would depend on the relative proportion of functional AB' to functional AC items. However, at least to this author's knowledge, there is as yet no data on negative transfer for AC relative to CD as a function of anticipation-interval length.

It was pointed out earlier in the paper that pilot work discovered that increasing the number of pairs in the list considerably amplified the initial positive transfer for AB' relative to CD. One explanation of this phenomenon is consistent with the mixed-list concept of the AB' paradigm. As the number of items per list is increased, the absolute number of items that can functionally serve as AB' items for the average subject is also increased. And as the number of functional AB' pairs increases, the likelihood of getting a greater initial absolute number of items correct would also increase. However, performance for CD should not increase correspondingly. If anything, initial absolute performance for CD should decrease as list length increases, since as list length increases, the retention interval also increases, placing a greater demand on memory.

The mixed-list (or mixed-paradigm) hypothesis may also explain another common phenomenon. In most transfer studies with the AB' paradigm, AB' performance in relation to CD exhibits either decreasing positive transfer, increasing negative transfer, or a combination of the two effects as a function of trials in list 2 learning. To explain this phenomenon we make four assumptions: (a) that a nominal AB' paradigm is functionally a mixed AB' and AC paradigm in which the proportions of AB' and AC items depend on the particular subject as well as the materials used, (b) that during list 2 learning the AB' items are learned more rapidly than the AC items, (c) that during list 2 learning the AB' items are learned more rapidly than the CD items, and (d) that during list 2 learning the CD items are learned more rapidly than the AC items. It can be readily seen that the last three assumptions are merely restatements of traditional transfer theory.

With these four assumptions we can define initial transfer for AB' relative to CD as the difference between the number of functional AB' items in list 2 and the mean initial-performance level for the CD control. If the number of functional AB' items is smaller than the average initial-performance level for CD, then initial negative transfer will occur for AB'. Once a subject has learned all the functional AB' items and begins to learn the functional AC items in the mixed list, the rate of increase in performance across trials for AB' will decrease in relation to the rate of increase in performance across trials for CD. And once the rate of increase in performance across trials for AB' falls below that of

CD, the typical interaction between AB' and CD in performance across trials will occur. Note that during the course of list 2 learning, an item that may initially have been a functional AC item can become a functional AB' item if the subject is able to resurrect the association learned in list 1 or if he eventually learns the association between the B' and B terms.

Unfortunately, although the mixed-paradigm hypothesis fairly successfully describes performance with the AB' paradigm, it does not easily lend itself to an experimental test of its own validity, since the relative proportions of AB' and AC items within a list are by definition subject-defined variables.

Notes

The present study was undertaken as partial fulfillment of the requirements for an undergraduate honors degree at the University of Iowa. The research was supported in part by Grants GB8392X and GY7251 from the National Science Foundation, the first a research grant to Professor R. W. Schulz for work on mediation, the second a grant in support of undergraduate research participation. The author acknowledges the aid and guidance of Professor Schulz, who was responsible for stimulating much of the thought and work that culminated in this paper. Revised manuscript received for publication April 7, 1971.

References

- Barnes, J. M., and Underwood, B. J. 1959. "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology* 58:97-105.
- Glaze, J. A. 1928. The association value of non-sense syllables. *Journal of Genetic Psychology* 35:255-269.
- Haagen, C. H. 1949. Synonymity, vividness, familiarity, and association value ratings of 400 pairs of common adjectives. *Journal of Psychology* 27:453-463.
- Postman, L. 1972. Transfer, interference, and forgetting. In *Experimental psychology*, ed. J. W. Kling and L. A. Riggs. In press.
- Richardson, J. 1967a. Transfer and the A-B anticipation interval in the A-B', A-B paradigm. *Journal of Verbal Learning and Verbal Behavior* 6:897-902.
- Richardson, J. 1967b. Latencies of implicit verbal responses and the effect of the anticipation interval on mediated transfer. *Journal of Verbal Learning and Verbal Behavior* 6:819-826.
- Richardson, J. 1968. Latencies of implicit associative responses and positive transfer in paired-associate learning. *Journal of Verbal Learning and Verbal Behavior* 7:638-646.
- Runquist, W. N., and Marshall, M. A. 1963. Transfer, synonymity, and antici-

- patory interval in paired-associate verbal learning. *American Journal of Psychology* 76:281-286.
- Schulz, R. W., and Lovelace, E. A. 1964. Mediation in verbal paired-associate learning: The role of temporal factors. *Psychonomic Science* 1:95-96.
- Schulz, R. W., and Weaver, G. E. 1968. The A-B, B-C, A-C mediation paradigm: The effects of variation in A-C study- and test-interval lengths and strength of A-B or B-C. *Journal of Experimental Psychology* 76:303-311.
- Spear, N. E., Mikulka, P. J., and Podd, M. 1966. Transfer as a function of time to mediate. *Journal of Experimental Psychology* 72:40-46.

CONTOUR-CONTINGENT COLOR AFTEREFFECTS: RETINAL AREA SPECIFICITY

C. F. Stromeyer III
Stanford University

Once an area of the retina was adapted to colored gratings, the aftereffect colors were judged to be most saturated when the test grating approximately coincided with the adapted area, less so as the overlap was reduced. The colors disappeared when the grating was moved just off the area and reappeared when the grating was moved very near it. In brief, the McCollough effect is highly area specific.

McCollough demonstrated a color aftereffect specific to the orientation of gratings of light and dark bars (1965). The aftereffect was built up by alternating a vertical grating of black stripes on one color with a horizontal grating on another color. Negative color aftereffects were seen on test gratings of black and white stripes, in a retinal orientation like that of the adaptation gratings.

The question in the present paper is whether the aftereffect is area specific. That is, does adapting a small area of the retina to colored gratings produce a color aftereffect that can be elicited by a small test grating placed almost anywhere on the retina, or must the grating fall on the originally exposed area? A nonlocalized aftereffect might reflect a learning process (an association between edge orientation and color) or, alternatively, an adaptation process involving cells with very large receptive fields. A localized aftereffect, on the other hand, might suggest that the effect depends on cells with small receptive fields—and would be consistent with the findings on many other sensory aftereffects.

Two classes of area-specific (localized) aftereffects may be distinguished. One class requires that the adaptation, or inducing, stimulus remain quite fixed on the retina. The afterimage is the typical case: a sharp afterimage is seen only when the inducing stimulus is not smeared over the retina. The figural aftereffect may also behave like an afterimage: the inducing stimulus produces a relatively fixed process on the retina, which distorts or displaces other contours subsequently placed near the affected area (Köhler and Wallach, 1944; Osgood and Heyer, 1952; Ganz, 1966).

A second class of area-specific aftereffects does not require that the inducing stimulus stay fixed on the retina. The subject can scan the stimulus, or the stimulus can be moved over the retina. However, these aftereffects are also, in a sense, area specific: they are restricted to the area of the retina exposed to the adaptation stimulus, or opposite aftereffects may be produced on adjacent areas of the retina. This class includes the short-term spiral aftereffect (Richards and Smith, 1969; Sekuler and Pantle, 1967), the 24-hour spiral aftereffect (Masland, 1969), the color aftereffect specific to the direction of edge motion (Stromeyer and Mansfield, 1970), and the increase in contrast threshold for a grating moving in the same direction as a high-contrast inspection grating (Pantle and Sekuler, 1969).

In addition to motion aftereffects, the second class of area-specific aftereffects also includes certain ones due to stationary contours. For example, adaptation to a grating both raises the contrast threshold for seeing a test grating of similar orientation and spatial frequency (Blakemore and Campbell, 1969) and causes an apparent frequency shift away from the adaptation frequency (Blakemore and Sutton, 1969; Blakemore, Nachmias, and Sutton, 1970). The frequency shift is area specific, for opposite effects can be produced on adjacent areas of the retina (Blakemore et al., 1970), and adaptation to a small-area grating may produce a frequency shift restricted to a small region of a large-area test grating (C. Blakemore, personal communication). The threshold elevation is also area specific, if we accept the argument that the two aftereffects are manifestations of a common neural mechanism (Blakemore and Nachmias, 1971).

Now, the McCollough effect does not behave like a striped afterimage, for Harris and Gibson (1968) have shown that the aftereffect can be built up by exposing the retina to randomly positioned colored gratings. Thus, the McCollough effect does not fall into the first class of area-specific effects. Does it fall into the second?

Several experiments have been performed to determine whether the aftereffect is restricted to the area of the retina exposed to the adaptation gratings. Murch (1969) built up aftereffects with two alternating colored gratings. The aftereffect colors appeared to extend over the entire surface of test patterns even twice the angular size of the gratings that had been fixated during adaptation. However, these subjects were instructed to ignore any color that moved with their eyes. Using Murch's technique, Stromeyer (unpublished findings) also found that after subjects had adapted to small-area color gratings, a large test grating appeared colored all over. But it is possible that small eye movements may

have generated a shallow color gradient here, which might have caused the color aftereffect to appear to spread beyond the adapted area; a shallow luminance (O'Brien, 1958) or color gradient (Lettvin, 1967; Krauskopf, 1967) can produce the perception of a homogeneous gray expanse or homogeneous color expanse.

Patterns with strong internal boundaries may help limit this spread. Harris (1969) has shown that opposite aftereffects can be built up on adjacent areas of the retina. The adaptation and test patterns he used were divided into four quadrants, each of which contained a vertical or horizontal grating measuring 4 deg on each side. The quadrants themselves were separated by .5-deg-wide black bars (C. S. Harris, personal communication). A green aftereffect specific to verticals was produced in one quadrant, and a red aftereffect, also specific to verticals, in an adjacent quadrant. Here, however, the black bars separating the quadrants and the fact that opposite effects were built up side by side may have prevented each aftereffect from spreading beyond its quadrant.

In the experiments that follow, the subjects were neither presented patterns with strong internal boundaries nor were they asked to judge an elusive border between a region of color and a region of no color on a large test pattern. Instead, they simply judged whether or not a small test grating appeared colored as it was moved into various retinal locations relative to the adapted area.

EXPERIMENT I

Method

—Subjects—Four Harvard students with normal color vision as tested with the AO H-R-R pseudoisochromatic plates served in the experiment. They were untrained.

—Apparatus—Figure 1 shows the adaptation and test display (with grating vertical). Vertical and horizontal gratings were projected alternately in the position shown in Figure 1. The black and white grating bars subtended 38 and 28 min of angle respectively, with luminances of 6 and 98 mL.

During adaptation the vertical grating was projected through a Wratten 55 filter (green, dominant λ with illuminant A, 524 nm; excitation purity, 62%; luminous transmittance, 25%) and the horizontal grating through a Wratten 26 filter (red, 621 nm; 100%; 19%). The same gratings, without colored filters, were used in the test condition. The projection screen (the large white rectangle shown in Figure 1) was placed at a 2-m viewing distance during the adaptation period and at various distances during the test. However, the gratings were always kept at a constant angular size and retinal spatial frequency, and the left edge of the gratings always fell on the

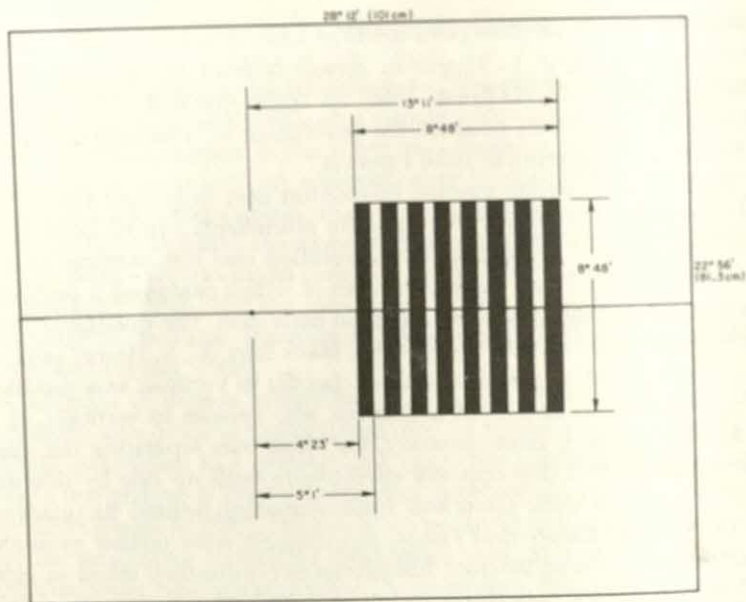


Fig. 1. Adaptation and test pattern (with vertical grating) at a 2-m viewing distance, Experiment I

vertical bisector of the screen. The fixation point was attached to a thin wire that could be moved horizontally (see Figure 1).

—Procedure—The subject sat 2 m from the screen, with his chin on a rest, and stared at the fixation point (the small dot in Figure 1) for a period of 20 min while the green-vertical and red-horizontal gratings were interchanged every 10 sec. Adaptation was followed by a 5-min rest. In the test, the gratings, without colored filters, were interchanged about every .5 sec. The fixation point was initially at the left edge of the screen, and the subject was told to continuously fixate the point. The experimenter then slowly moved the point horizontally until the subject indicated that the aftereffect appeared maximally saturated; the subject could ask the experimenter to move the point left or right. One judgment was made with the screen at 2 m, and then one each at 3 and 4 m, to see whether the aftereffect obeys Emmert's law.

Results

All four subjects reported that the vertical test grating was tinted red and the horizontal green when the gratings overlapped the area of the retinas originally exposed to the colored gratings. A flash of color was seen each time the gratings were interchanged. Table 1 shows the angular distance between the fixation point and the left edge of the grat-

ings at which the aftereffects appeared most saturated. On the average, the fixation point was set about 5 deg to the left of the gratings at all viewing distances. During adaptation the distance had been 4 deg 23 min. Thus, the aftereffect appeared most saturated when the retinal locations of the adaptation and test gratings approximately coincided.

Table 1. Angle in deg and min between fixation point and left edge of test gratings at which aftereffects appeared maximally saturated

Subject	Viewing distance		
	2 m	3 m	4 m
RM	5°3'	4°56'	5°3'
DH	5°14'	6°29'	5°27'
AF	4°50'	3°37'	4°43'
SC	4°47'	5°32'	5°29'
Geometric means	4°57'	5°2'	5°10'

EXPERIMENT II

This experiment examined the degree of overlap between the retinal locations of the adaptation and test gratings that would be required to see any aftereffect at all. Presumably, a highly area-specific aftereffect would disappear when the test grating was moved just off the adapted area.

Method

—Subjects—The three subjects, including the author, were highly experienced with the McCollough effect but had never made the judgments required in the present experiment.

—Apparatus—For adaptation, vertical and horizontal gratings with .2-deg black and white bars were rear-projected on a ground-glass screen by projectors with 500-w bulbs. Each grating was 7 deg 8 min square at a viewing distance of 82 cm. The luminance of black and white bars was 36 and 600 mL.

The vertical grating was projected through a Wratten 53 filter (green, dominant λ with illuminant A, 552 nm; excitation purity, 82%; luminous transmittance, 8%) and the horizontal grating through a Wratten 34a filter (magenta, 574c nm; 92%; 4%). (These filters produce remarkably strong aftereffects, as evidenced by many observers in this and other experiments.) A tiny red fixation light was placed either at the center of the left edge of the gratings or at the center of the top edge. The gratings were interchanged every 1 sec. On each presentation the gratings were projected in the same position or shifted 180 deg in phase relative to the fixation point, according to a random schedule. This technique was used to assist in homogeneously

adapting a given retinal area to black and colored bars (Harris and Gibson, 1968).

The test pattern, illustrated in Figure 2, was composed of two gratings that formed a square of 6 deg 44 min, from a viewing distance of 82 cm. The retinal width of the bars matched the adaptation gratings. The pattern was illuminated by fluorescent lights such that the luminance of the white bars was 20 mL. Either the pattern was oriented as shown in Figure 2, in which case it could be moved horizontally relative to a stationary fixation point on the horizontal bisector of the pattern, or was oriented 90 deg clockwise from the position in Figure 2 and could be moved vertically relative to a fixation point on the vertical bisector of the pattern. The pattern was moved along a track by means of strings. A white cardboard surround was placed behind the test pattern.

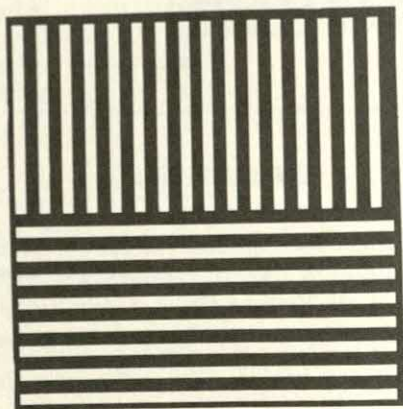


Fig. 2. Test pattern, Experiment II

—Procedure—To build up an aftereffect in each session, the subject, his head on a chin rest, stared at the small fixation light for three 5-min periods of exposure to the colored gratings, each exposure separated by a short rest. In the first test session the fixation light was positioned at the left edge of the gratings (to adapt an area to the right of the fixation point), and the subject moved the test pattern *left past* the fixation point to the position at which all color just disappeared from the gratings, and then *right toward* the fixation point until the first trace of color appeared. The pattern was moved in one direction at a time, and no opposite movement was permitted. Each judgment was repeated five times. In the second test session, the fixation light was positioned at the middle of the top edge (to adapt an area below the fixation point), and the subject moved the pattern *up past* the fixation point to the position at which all color just disappeared from the gratings, and then *down toward* the fixation point until the first trace of color appeared. Each of these judgments was repeated three times.

Results

Table 2 shows the results. The test pattern appeared colorless when positioned 40.4 min (geometric mean for all subjects) to the left of the adapted area (first session), or when positioned 27.6 min above the adapted area (second session). The first trace of color appeared when the pattern was positioned 20.0 min to the left of the adapted area, or when positioned 8.5 min above the adapted area.¹ The judgments of each subject were highly consistent, as shown by the small standard error of the means (σ_M).

Table 2. Geometric means of angle in min between fixation point and right (or bottom) edge of test gratings at which all color just disappeared from or color just appeared on the grating

Subject	Color disappeared		Color appeared	
	Grating moved left (up)	σ_M	Grating moved right (down)	σ_M
First session				
MW	-41.7	7.1	-21.9	7.6
CS	-34.4	7.5	-18.1	.9
BJ	-44.7	4.0	+35.8	4.2
Group means	-40.4		-20.0†	
Second session				
MW	(-31.5)	(1.5)	(-19.1)	(1.1)
CS	(-23.2)	(2.6)	(+ 2.2)	(5.9)
Group means	(-27.6)		(- 8.5)	

Note: Positive angles indicate direction right (down) from the fixation point; σ_M is the standard error of the mean based on five (three) judgments.

† Mean excludes subject BJ.

DISCUSSION

Experiment I demonstrated that the aftereffect was most saturated when the test pattern approximately overlapped the adapted area, whereas Experiment II demonstrated that the aftereffect disappeared when the test pattern was moved only about .5 deg off the adapted area, and then reappeared when the pattern was brought within about .3 deg of the adapted area. In Experiment II, all subjects reported that the aftereffect appeared most saturated when the test pattern largely overlapped the adapted area and that the aftereffect desaturated when

the extent of overlap was reduced. In spite of the weakness of the colors when the test pattern was moved just off the adapted area, the subjects' judgments were highly consistent. When the pattern partially overlapped the adapted area, no subject reported seeing a sharp border between an area of color and an area of no color on the test pattern, although the aftereffect appeared more saturated on the adapted side.

Thus, we may conclude that the McCollough aftereffect is area specific in the sense that the effect is seen only when the test pattern overlaps the retinal area originally exposed to colored gratings. However, in some cases, with a large test pattern, subjects may be unable to report that the aftereffect color is strictly confined to the adapted region.

Notes

Received for publication June 4, 1971.

1. The value of 20.0 min is based on the data for subjects CS and MW alone. Subject BJ required considerably more overlap of the test pattern and adapted area to see the first trace of color than did the other two subjects. Later questioning revealed that BJ had set the pattern so that both the green and the pink aftereffects were seen, whereas subjects CS and MW had set it so that a trace of either aftereffect was seen.

References

- Blakemore, C., and Campbell, F. W. 1969. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology* 203:237-260.
- Blakemore, C., and Nachmias, J. 1971. The orientation specificity of two visual after-effects. *Journal of Physiology* 23:157-174.
- Blakemore, C., Nachmias, J., and Sutton, P. 1970. The perceived spatial frequency shift: Evidence for frequency-selective neurones in the human brain. *Journal of Physiology* 210:727-750.
- Blakemore, Colin, and Sutton, Peter. 1969. Size adaptation: A new aftereffect. *Science* 166:245-247.
- Ganz, Leo. 1966. Mechanism of the figural aftereffects. *Psychological Review* 73:1506-1507.
- Harris, C. S. 1969. Retinal localization of orientation-specific color after-effects. *Journal of the Optical Society of America* 59:504.
- Harris, C. S., and Gibson, A. R. 1968. Is orientation-specific color adaptation in human vision due to edge detectors, afterimages, or "dipoles"? *Science* 162:162-63.
- Köhler, W., and Wallach, H. 1944. Figural after-effects: An investigation of visual processes. *Proceedings of the American Philosophical Society* 88: 269-357.

- Krauskopf, J. 1967. Heterochromatic stabilized images: A classroom demonstration. *American Journal of Psychology* 80:634-637.
- Lettvin, J. Y. 1967. The colors of colored things. *Quarterly Progress Report of the Research Laboratory of Electronics, Massachusetts Institute of Technology* 87:193-229.
- McCollough, C. 1965. Color adaptation of edge-detectors in the human visual system. *Science* 149:1115-1116.
- Masland, Richard H. 1969. Visual motion perception: Experimental modification. *Science* 165:819-821.
- Murch, Gerald M. 1969. Size judgments of McCollough afterimages. *Journal of Experimental Psychology* 81:44-48.
- O'Brien, V. 1958. Contour perception, illusion and reality. *Journal of the Optical Society of America* 48:112-119.
- Osgood, C. E., and Heyer, A. W. 1952. A new interpretation of the figural after-effect. *Psychological Review* 59:98-118.
- Pantle, Allan, and Sekuler, Robert. 1969. Contrast response of human visual mechanisms sensitive to orientation and direction of motion. *Vision Research* 9:397-406.
- Richards, Whitman, and Smith, Robert A. 1969. Midbrain as a site for the motion after-effect. *Nature* 223:533-534.
- Sekuler, Robert, and Pantle, Allan. 1967. A model for aftereffects of seen motion. *Vision Research* 7:427-439.
- Stromeyer, C. F. III, and Mansfield, R. J. W. 1970. Colored aftereffects produced with moving edges. *Perception and Psychophysics* 7:108-114.

continued from page 188

- Kapleau, Philip, ed. *The Wheel of Death: A Collection of Writings from Zen Buddhist and Other Sources on Death, Rebirth, Dying*. New York: Harper & Row, 1971. Pp. xviii, 110. \$4.95.
- Kling, J. W., and Riggs, Lorrin A., eds. *Woodworth and Schlosberg's Experimental Psychology*. 3d ed. New York: Holt, Rinehart and Winston, 1971. Pp. xiv, 1,279.
- Krishna, Gopi. *The Biological Basis of Religion and Genius*. New York: NC Press, 1971. Pp. ix, 64. \$4.50; paperback, \$1.95.
- Laing, R. D. *"The Politics of the Family" and Other Essays*. New York: Pantheon Books, 1971. Pp. ii, 133. \$4.95.
- Laing, R. D., and Cooper, David. *Reason and Violence: A Decade of Sartre's Philosophy, 1950-1960*. New York: Pantheon Books, 1971. Pp. 184. \$5.95.
- Landfield, A. W. *Personal Construct Systems in Psychotherapy*. Chicago: Rand McNally, 1971. Pp. ix, 220.
- Landis, Bernard, and Tauber, Edward S., eds. *In the Name of Life: Essays in Honor of Erich Fromm*. New York: Holt, Rinehart and Winston, 1971. Pp. xiv, 333. \$10.95.
- Lannoy, Richard. *The Speaking Tree: A Study of Indian Culture and Society*. New York: Oxford University Press, 1971. Pp. xxvi, 466. \$15.00.
- Lee, Wayne. *Decision Theory and Human Behavior*. New York: Wiley, 1971. Pp. x, 352. \$13.95.
- Leedy, Jack J., ed. *Compensation in Psychiatric Disability and Rehabilitation*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xx, 361. \$15.00.
- MacDermot, Violet. *The Cult of the Seer in the Ancient Middle East: A Contribution to Current Research on Hallucinations Drawn from Coptic and Other Texts*. Berkeley and Los Angeles: University of California Press, 1971. Pp. xii, 829. \$24.00.
- MacDonald, Angus. *Middle Ground*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. 375. \$7.95.
- McGaugh, James L., ed. *Psychobiology*. New York: Academic Press, 1971. Pp. xiii, 366. \$14.50.
- McReynolds, Paul, ed. *Advances in Psychological Assessment*. Vol. 2. Palo Alto, Calif.: Science and Behavior Books, 1971. Pp. 395. \$11.50.
- Madow, Leo, and Snow, Laurence H., eds. *The Psychodynamic Implications of the Physiological Studies on Psychomimetic Drugs*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xiii, 81. \$7.95.
- Milio, Nancy. *9226 Kercheval: The Storefront That Did Not Burn*. Ann Arbor: University of Michigan Press, 1971. Pp. 209. Paperback, \$2.45.
- Mischel, Theodore, ed. *Cognitive Development and Epistemology*. New York: Academic Press, 1971. Pp. xv, 423. \$16.40.
- Otto, Herbert S., ed. *The New Sexuality*. Palo Alto, Calif.: Science and Behavior Books, 1971. Pp. xiii, 289. \$8.95.
- Patterson, Gerald R., and Gullion, Elizabeth M. *Living with Children: New Methods for Parents and Teachers*. Champaign, Ill.: Research Press, 1971. Pp. iv, 117.
- Pribram, K. H. *Languages of the Brain: Experimental Paradoxes and Principles in Neuropsychology*. Englewood Cliffs. N. J.: Prentice-Hall, 1971. \$9.50.

continued on page 248

FREE RECALL OF ITEMS PRESENTED AFTER MASSED- AND DISTRIBUTED-PRACTICE ITEMS

David G. Elmes, William I. Greener, and William C. Wilkinson

Washington and Lee University

When critical items were presented after items receiving either massed or (at repetition lags of 3 or 10 events) distributed practice, free recall was better on the distributed- than massed-practice items, but worse on the critical items after distributed- than after massed-practice items. Differential encoding, rather than consolidation, better accounts for the differences with the two kinds of practice.

Several recent studies (see Melton, 1970) have demonstrated that the retention of items receiving distributed practice exceeds the retention of items receiving massed practice. Two general kinds of explanations have been offered to account for this effect (Bjork and Allen, 1970). The first, the *consolidation hypothesis*, asserts that the longer the interval between repetitions, the greater the consolidation (strength) of the item in memory. The second class of explanations, the *differential-encoding hypotheses*, states that items are encoded better in distributed than in massed practice, presumably because subjects do not encode the second presentation of an item adequately in massed practice (Underwood, 1970) or because more new information is encoded during the second presentation of an item in distributed than in massed practice (Melton, 1970).

Bjork and Allen (1970) have provided evidence in favor of differential encoding. Their study involved variation of task difficulty between item repetitions rather than distribution of practice. Contrary to the consolidation hypothesis, Bjork and Allen found slightly better retention when a difficult task separated the two item presentations than when an easy task intervened between repetitions.

The present studies tested the differential-encoding and consolidation hypotheses with a manipulation of massed versus distributed practice. Free recall of 'critical' items presented after those practice items was examined. If more adequate processing of items occurs in distributed than massed practice, as supposed by the notion of differential encoding, then recall of critical items presented after distributed-practice

items should be poorer than recall of critical items presented after massed-practice items. On the other hand, if time in storage is the important variable, as the consolidation hypothesis holds, then recall of critical items should be unaffected by the type of practice on the previous item in the list.

METHOD

—Experiment I—Thirty Washington and Lee University students in introductory psychology heard and free-recalled two lists of 16 two-syllable nouns that had Thorndike and Lorge (1944) frequencies of at least 50 per 1,000,000 words.

A sample list appears in Table 1, showing the massed- and distributed-prac-

Table 1. Portion of list illustrating task and materials, Experiment I

	weather	
	anger	
	midnight	
	uncle	
<i>massed practice</i>	{ circle	
	{ circle	
	purpose	<i>critical after massed practice</i>
	empire	
<i>massed practice</i>	{ number	
	{ number	
<i>distributed practice</i>	purpose	<i>critical after massed practice</i>
<i>distributed practice</i>	honey	<i>critical after distributed practice</i>
	language	
	language	
	empire	
	honey	<i>critical after distributed practice</i>
	:	
	midnight	
	weather	
	uncle	
	anger	

tice items of interest as well as the critical items presented thereafter. Several features of the lists are notable: (a) all items were presented twice, (b) buffer items began and ended the lists to minimize primacy and recency effects on the items of interest, (c) a given critical item (say, "purpose" or "honey") always occurred after either a massed-practice item or after a distributed-practice item, (d) all criticals had a repetition lag of 3 events, and (e) each list contained two critical items after massed-practice items and two critical items after distributed-practice items, with one of the distributed items having a repetition lag of 3.

A tape recorder presented the words at a 3-sec rate, and subjects had 60 sec for written free recall. Three word orders of the two lists were formed. Ten subjects learned one set of these forms. Counterbalancing assured that each list form occurred first equally often.

—Experiment II—This experiment differed from the first one as follows: (a) each subject free-recalled two 44-item lists, (b) each list contained three critical items after the massed-practice items, three criticals after distributed-practice items with a repetition lag of 3, and three criticals after distributed-practice items with a repetition lag of 10, (c) 90 sec was allowed for written free recall, (d) critical items had a repetition lag of 13 events, and (e) ten introductory psychology students at Washington and Lee University served as subjects.

RESULTS AND DISCUSSION

The percentages of massed- and distributed-practice items recalled in both lists of Experiments I and II are shown in the upper portion of Table 2. Consistent with previous results (Melton, 1970), distrib-

Table 2. Percentage correct free recall on and after massed-practice (MP) and distributed-practice (DP) items

Experiment	Repetition lag		
	MP 0	DP 3	DP 10
Practice items			
I	33	42	
II	12	21	23
Critical items			
I	49	35	
II	38	25	18

uted practice resulted in better free recall than did massed practice [$F(1, 29) = 5.26, p < .05$ in Experiment I; $F(2, 18) = 5.33, p < .025$ in Experiment II]. Neither lists nor the interaction of lists and repetition lag was a significant source of variance ($F_s \leq 1.27$). In Experiment II the free recall under distributed practice was not reliably better with a repetition lag of ten than with one of three, by the Tukey (a) test (Winer, 1962).

Recall of critical items presented after massed- and distributed-practice items is of primary interest. As shown in the lower portion of Table 2, recall was lower for critical items presented after distributed-

practice items than when the critical items occurred after massed-practice items. This effect was reliable in both studies [$F(1, 29) = 5.22$, $p < .05$ in Experiment I; $F(2, 18) = 4.67$, $p < .025$ in Experiment II]. No interactions approached conventional significance levels ($F_s < 1$).

Lower recall of critical items presented after distributed-practice items than after massed-practice items does not seem readily interpretable in terms of consolidation. Rather, the superior recall of critical items presented after massed-practice items indicates that subjects operate on massed- and distributed-practice items in different ways. Two related ways of differential processing could account for the present results. In the first place, if subjects ignore or pay less attention to the second presentation of an item in massed than in distributed practice, the critical item immediately following a massed-practice item should be learned more easily than a critical item following a distributed-practice item. Secondly, if more new storage and retrieval cues accrue to distributed- than to massed-practice items, it seems reasonable to expect attenuated performance on items after distributed-practice ones. Although evidence is lacking in the present studies to enable a choice between these processes, it may be that both factors operate to produce the different practice effects.

Notes

This research was supported by a Robert E. Lee Research Grant. Portions of these data were presented by the second author at the meetings of the Virginia Academy of Science, Blacksburg, May, 1971. Requests for reprints should be sent to David G. Elmes, Department of Psychology, Washington and Lee University, Lexington, Virginia 24450. Received for publication June 3, 1971.

References

- Bjork, R. A., and Allen, T. W. 1970. The spacing effect: Consolidation or differential encoding? *Journal of Verbal Learning and Verbal Behavior* 9:567-572.
- Melton, A. W. 1970. The situation with respect to the spacing of repetitions and memory. *Journal of Verbal Learning and Verbal Behavior* 9:596-606.
- Thorndike, E. L., and Lorge, I. 1944. *The teacher's word book of 30,000 words*. New York: Columbia University Teachers College.
- Underwood, B. J. 1970. A breakdown of the total-time law in free-recall learning. *Journal of Verbal Learning and Verbal Behavior* 9:573-580.
- Winer, B. J. 1962. *Statistical principles in experimental design*. New York: McGraw-Hill.

THE McCOLLOUGH EFFECT CREATED BY COMPLEMENTARY AFTERIMAGES

Gerald M. Murch and Joy Hirsch

Portland State University

A new version of the McCollough effect was found in three experiments in which line orientation and color were separated. The results indicate that a complementary afterimage evoked by a lined pattern can serve as an effective stimulus to create the McCollough effect—which suggests that the effect is not adequately explained by adaptation of color-coded edge detectors.

McCollough (1965) was the first to observe the visual effect that a pattern of vertical stripes on a red ground, alternating with a pattern of horizontal stripes on a green ground, has upon the subsequent perception of a test pattern of vertical and horizontal lines on a white ground. Observers usually report an apparent reddish color on the horizontal and a greenish color on the vertical portions of the neutral test pattern. This is known as the *McCollough effect*.

Based on Hubel and Wiesel's observation (1959, 1968) that the visual system of the vertebrate contains edge detectors sensitive to line orientation, McCollough's model (1965) to explain the effect assumed that these detectors were color sensitive and subject to fatigue after successive stimulation, responding with decreased sensitivity to those wavelengths with which they had been most recently stimulated. According to McCollough, the effect was the result of a fatiguing of the vertical-line detectors sensitive to red, a fatiguing such that the test pattern's vertical lines were then processed by the nonadapted vertical receptors for green and blue. Similarly, prolonged viewing of the horizontal pattern on green fatigued the horizontal-line detectors sensitive to green, so that the test pattern's horizontal lines appeared reddish.

Stromeyer (1969) further observed that alternating a single pattern on a colored background with another, orthogonally oriented pattern on a neutral ground would also produce the usual McCollough effect. An inspection pattern of vertical lines on a red ground ('red verticals') alternated with an inspection pattern of horizontal lines on a white ground produced a greenish aftereffect on the verticals and a reddish aftereffect on the horizontals of a test pattern. Stromeyer explained his

results by hypothesizing that the color-sensitive vertical and horizontal line detectors are organized in an opponent manner, since the red verticals had produced both a green aftereffect on the verticals and a red aftereffect on the horizontals of his neutral test pattern.

Yet, an assumption of opponently organized line detectors is not required to explain Stromeier's results. It could also be that the prolonged viewing of the red verticals in his experiment fatigued the retinal receptors sensitive to red, so that the white background of the horizontal inspection pattern was processed primarily by retinal units specific for blue and green. Over several alternations the *effective* stimulus pattern would thus be red verticals and blue-green horizontals. This explanation assumes that a complementary afterimage paired with a lined pattern can be an effective stimulus for the McCollough effect.

EXPERIMENT I

This experiment sought to test the possibility that a complementary afterimage can induce the McCollough effect. The color stimulus and the lined pattern were separated, in such a way that a colored field was alternated with a lined pattern on a white ground during inspection. If the complementary-afterimage hypothesis holds, then red alternating with, say, verticals should result in a red aftereffect of the verticals on a neutral test pattern, as the effective stimulus is the red complementary afterimage (blue-green) to the first inspection pattern, plus the verticals on the other inspection pattern.

Method

—Apparatus—The two inspection patterns (unlined colored field and vertical or horizontal lines) were alternately projected on the same portion of a white screen via Kodak Carousel projectors. Either a Wratten 61 (green) or 25a (red) filter provided the unlined colored field. Neutral density filters were also mounted with the colored filters such that the intensity of each projection measured 10.0 ftc at the screen. Both inspection patterns subtended 11.6 deg of visual angle; each black vertical or horizontal line subtended 12.4 min of arc, with the white area between the lines subtending 24.9 min of arc. The frequency of presentation, the duration, and the interval between presentations were all controlled by solid-state programming equipment.

Two more projectors provided the test patterns, vertical or horizontal lines on a white background. The dimensions and intensity of each test pattern were the same as for the inspection patterns. The two test patterns were projected on the screen simultaneously and side by side, the space between them subtending 35 min of arc.

—Procedure—Each of the 31 subjects was run individually. Upon reporting for the experiment the subject was seated behind a control panel and was asked to look through a binocular viewer at the screen 10 ft away. He was informed that he would see a series of projections on the screen and that he was to look carefully at each projection but not to maintain a strict fixation point. After the last projection, he was told that he would probably see a 'blob' of color floating in space and moving about as he shifted his gaze. He was to wait until this color disappeared and then turn on the test patterns by using the two switches on the control panel in front of him.¹ He was then asked to study the test patterns and to name any color he saw between the dark lines of each of the two patterns. Also, he was required to match that color on a projection colorimeter (similar in design to Riggs', 1964).

Of the subjects 16 got the green stimulus alternating ten times with the horizontal pattern ('green *plus* horizontals'). The others got the red field alternating with the vertical pattern ('red *plus* verticals'). Each stimulus pattern was presented for 10 sec with a 1-sec interval between patterns.

After the subject had indicated the colors for the vertical and horizontal test patterns and matched them with the colorimeter, a disinhibition series was begun. The inspection color was paired with the other line pattern (if the pair green plus verticals was used in inspection, green plus horizontals was used in disinhibition), and a single alternation of the new inspection patterns was presented. After the 'free-floating' afterimage disappeared, the subject turned on the test patterns and named the color on each. This procedure of a single alternation of the orthogonal line orientation and color followed by the test patterns was repeated until the subject indicated that the test patterns appeared colorless.

Results

Of those subjects who were first presented a red field alternating with vertical lines, five reported that the horizontals on the test pattern appeared green, the verticals colorless. Two reported the other half of the effect: the verticals on the test pattern appeared red. Two saw both halves of the effect (red aftereffect on verticals and green on horizontals), and six reported no aftereffect. Of those subjects who were first shown a green field alternating with horizontals, ten subjects reported the red aftereffect on the verticals, and only one saw the green aftereffect on the horizontals of the test pattern. One reported the entire effect, and four reported no aftereffects.

The colorimetric matches were translated into CIE x,y coordinates. While these matches were intended primarily as a control for the accuracy of color names, they also were used in several instances by subjects who indicated no McCollough effect by color name but then selected an appropriate match on the colorimeter.

The conditions of this experiment produced a color aftereffect that differs from the usual McCollough effect. Instead of the aftereffect be-

ing one in which the inspection of red verticals produces at the test the approximate inspection-color complement (green) on the inspection orientation (vertical), here the inspection of red plus verticals produced the inspection color itself (red) on the inspection orientation (vertical) or the opponent color (green) on the orthogonal orientation (horizontal). But if one considers the inspection patterns in terms of a complementary afterimage evoked by the lined pattern on white, then the results of this experiment do conform to previous work on the McCollough effect, for it was the complement of the *afterimage* of the inspection color that was seen on the inspection orientation during testing (or the other half of the effect, the inspection color on the orthogonal orientation). These observations do not refute Stromeyer's postulation of opponently organized edge detectors but do support the hypothesis that a complementary afterimage may serve as a stimulus for an orientation-specific color aftereffect.

EXPERIMENT II

If in fact the complementary afterimage evoked by the lines on the neutral inspection pattern is responsible for the reversed McCollough effect, then the effect should also occur if both conditions of Experiment I are combined in such a way that both red and green fields and vertical and horizontal lines are alternately presented. The complementary afterimages of those fields should occur on, or be evoked by, the following presentation of patterns of vertical and horizontal lines, thus producing a reversed McCollough effect on the test patterns.

Method

A total of 28 subjects observed this inspection sequence alternate ten times: unlined green field, vertical lines on white, unlined red field, horizontal lines on white. Each stimulus element was projected on the screen for 10 sec with a 1-sec interval between elements. In all other details the study paralleled Experiment I.

Results

The frequency of color names and the CIE x, y coordinates generated by all 28 subjects are shown in Table 1. A little less than half the subjects reported the expected greenish color on the vertical test pattern and the reddish color on the horizontal test pattern. Six subjects saw only the red half of the effect; one reported only the green half. The other subjects reported no color aftereffect. Again, the resultant effects—here,

Table 1. Frequency of color names and associated mean CIE x, y coordinates for vertical and horizontal test patterns, Experiment II

Color name		N	Color match			
Vertical	Horizontal		Vertical		Horizontal	
			x	y	x	y
Green	Red	13	.2312	.5569	.2541	.5425
	Red	6	.2305	.5335	.2367	.5297
Green		1	.2270	.5470	.2301	.5400
No aftereffect		8	.2318	.5400	.2292	.5340

green on the test verticals and red on the test horizontals after inspection of red plus horizontals and green plus verticals—must result from complementary afterimages processed in conjunction with specific line orientations.

EXPERIMENT III

One way to test the influence of the adaptive state of the retina at the time the inspection lines are presented would be to lengthen the interval between stimulus elements. If a green field is alternated with horizontal lines on a white ground, for example, lengthening the interval between the presentation of the green and the horizontals may be assumed to allow the retinal units to return to base-level functioning prior to the introduction of those horizontals. The McCollough effect occurs with intervals between the alternating patterns as long as 60 sec (Hajos, 1969), but if the effective stimulus is the *afterimage* of the colored field, increases in the interval between field and horizontals should result in a reduction in the reported frequency of the effect.

Method

This experiment followed the design of Experiment I, with the exception that four interstimulus intervals of 1, 5, 20 and 50 sec were employed. Sixteen subjects were observed for all interstimulus intervals. The order of intervals was randomized across subjects. By using the technique of disinhibition, we could observe all subjects at all intervals in a single one-hour sitting. To disinhibit the effect, the line orientation was changed to vertical and a single alternation of green field and vertical lines was presented, followed by the two test patterns. This procedure was continued until the subject described both test patterns as colorless.

All observers were shown a solid green field alternating with a pattern of horizontal lines. After ten alternations the subject turned on one test pattern of vertical and another of horizontal lines and indicated by name and colorimeter rating any color on the test patterns.

Results

The results are shown in Table 2, which indicates color names for the vertical and horizontal test patterns as a function of interstimulus interval between inspection patterns. (The matches with the colorimeter again controlled for the accuracy of color names.) After inspection of green horizontals, there was, as close examination of Table 2 shows, a *decrease* in the number of red aftereffects reported on the vertical test pattern alone (second row) as the interstimulus interval of inspection increased, and a progressive *increase* in red aftereffects reported on both vertical and horizontal test patterns (fifth row).

Table 2. Frequency of color names as a function of interstimulus interval for vertical and horizontal test patterns, Experiment III

Reported color name		Frequency of reported name			
Vertical	Horizontal	1 sec	5 sec	20 sec	50 sec
Red	Green	1	2	1	2
Red		10	5	5	2
	Green	1	1	0	0
	Red	0	0	3	2
Red	Red	0	6	7	9
No aftereffect		4	2	0	1

These results suggest that as the interstimulus interval was increased, the state of adaptation of the retinal color receptors for green on the horizontals tended to return to neutral, producing a lesser red response to those inspection horizontals with the receptors' adaptation for green. Thus, the effective afterimage stimulus was no longer processed with horizontal lines and, assuming that the green field was then processed equally by receptors specific to all orientations, the dominant aftereffect became one of red on both test patterns.

DISCUSSION

The observation that a complementary afterimage can serve as the effective stimulus for the McCollough effect when the afterimage is produced against a background pattern of vertical or horizontal lines has some interesting ramifications for the theoretical interpretations of the McCollough effect. The adaptation of color-coded edge detectors no longer appears to be an adequate explanation. The more recent work of Hubel and Wiesel (1968) indicates that edge detectors in the cor-

tex of the primate demonstrate little selectivity for wavelength, while cells in the lateral geniculate body demonstrate primarily opponent-color properties. As has been suggested earlier (Hirsch and Murch, 1971), color adaptation may occur at the lateral geniculate body in conjunction with cortically located vertical and horizontal detectors. However, extensive behavioral and physiological observation will be required before the mechanism mediating the McCollough effect is completely understood.

Notes

This research was supported by Research Grant NSF GB11817 from the National Science Foundation, Division of Psychobiology. The authors thank Barbara Breck for her aid in data collection. Received for publication April 15, 1971; revision received June 22, 1971.

1. This procedure allowed a subject to familiarize himself with a complementary afterimage in order that he might distinguish it from a McCollough effect.

References

- Hajos, A. 1969. Verlauf formenspezifischer Farbadaptationem im visuellen System des Menschen. *Psychologische Beiträge* 11:95-114.
- Hirsch, J., and Murch, G. M. 1971. A new version of the McCollough effect. Presented at the meeting of the Western Psychological Association, San Francisco, April.
- Hubel, D. H., and Wiesel, T. N. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology* 148:574-591.
- Hubel, D. H., and Wiesel, T. N. 1968. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Receptive fields and functional architecture of the monkey striate cortex. *Journal of Physiology* 160:106-154. 195:215-243.
- McCollough, C. 1965. Color adaptation of edge-detectors in the human visual system. *Science* 149:115-116.
- Riggs, L. A. 1964. The projection color mixer. *American Journal of Psychology* 77:129-134.
- Stromeyer, C. F. 1969. Further studies of the McCollough effect. *Perception and Psychophysics* 6:105-110.

continued from page 236

- Ransom, Arthur J. *An Introduction to Social Psychiatry*. Baltimore: Penguin Books, 1971. Pp. 168.
- Roberts, H. J. *The Causes, Ecology, and Prevention of Traffic Accidents: With Emphasis upon Traffic Medicine, Epidemiology, Sociology, and Logistics*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xxii, 1016. \$39.50.
- Rodman, Hyman. *Lower-class Families*. New York: Oxford University Press, 1971. \$8.00.
- Ryle, Anthony. *Neurosis in the Ordinary Family: A Psychiatric Survey*. New York: Barnes & Noble, 1971. \$2.50.
- Savage, C. Wade. *The Measurement of Sensation: A Critique of Perceptual Psychophysics*. Berkeley and Los Angeles: University of California Press, 1970. Pp. 578. \$15.00.
- Schachter, Stanley. *Emotion, Obesity, and Crime*. New York: Academic Press, 1971. Pp. vii, 195. \$7.50.
- Schmaltz, Leonard W., ed. *Scientific Psychology and Social Concern*. New York: Harper & Row, 1971. Pp. x, 502. Paperback, \$6.75.
- Schrier, Allan M., and Stollnitz, Fred, eds. *Behavior of Nonhuman Primates: Modern Research Trends*. New York: Academic Press, 1971. Pp. xv, 206. \$9.50.
- Seigenthaler, John, ed. *A Search for Justice*. Nashville: Aurora Publishers, 1971. Pp. x, 416. \$10.95.
- Simonson, Ernst, ed. *Physiology of Work Capacity and Fatigue*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xxvi, 571. \$33.50.
- Slavson, S. R., ed. *The Fields of Group Psychotherapy*. New York: Schocken Books, 1971. Pp. xxix, 338. Paperback, \$3.95.
- Snyder, Solomon H. *Uses of Marijuana*. New York: Oxford University Press, 1971. Pp. 127. \$5.95.
- Stuart, Richard B. *Trick or Treatment: How and When Psychotherapy Fails*. Champaign, Ill.: Research Press, 1971. Pp. ii, 201.
- Tart, Charles T. *On Being Stoned: A Psychological Study of Marijuana Intoxication*. Palo Alto, Calif.: Science and Behavior Books, 1971. Pp. xvii, 333. \$7.95.
- Tiger, Lionel, and Fox, Robin. *The Imperial Animal*. New York: Holt, Rinehart and Winston, 1971. Pp. xi, 308. \$6.95.
- Varela, Jacobo A. *Psychological Solutions to Social Problems: An Introduction to Social Technology*. New York: Academic Press, 1971. Pp. xiv, 317. \$12.50.
- Vernon, M. D. *Reading and Its Difficulties: A Psychological Study*. New York: Cambridge University Press, 1971. \$10.50.
- Vernon, Philip E. *The Structure of Human Abilities*. New York: Barnes & Noble, 1971. \$3.00.
- Viorst, Judith. *The Tenth Good Thing about Barney*. New York: Atheneum Publishers, 1971. Pp. 24. \$3.95.
- Vygotsky, Semenovich Lev. *The Psychology of Art*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xi, 305. \$12.50.
- Walcher, Dwain. *Early Childhood*. New York: Academic Press, 1971. \$11.00.
- Whaley, Donald L., and Malott, Richard W. *Elementary Principles of Behavior*. New York: Appleton-Century-Crofts, 1971. Pp. ix, 454. Paperback, \$6.95.

continued on page 260

A DEMONSTRATION OF COMPLEX RULE LEARNING IN CHOICE PREDICTION

Mari R. Jones and James R. Erickson

Ohio State University

Four groups ($N = 16$) of college students predicting 312 trials of a complex four-choice test sequence all learned to respond to rule-determined rather than event-run cues; accuracy increased with training on similar rules in different run-length distributions. This ability to abstract and discriminate complex relationships within events regardless of event-run properties implies a rule-learning rather than run-theory view of sequential learning.

On the basis of experiments with artificial grammars, Braine (1963) has argued that what is learned in an utterance is the temporal location of elements in a large hierarchical scheme. The implications of this approach for sequential learning are rather different than those of either run theory (Myers, 1970) or rule theory (Restle, 1967; Restle and Brown, 1970).

Run theory, for example, has considered subjects' sensitivity to the temporal order of events, but these have been *homogeneous* events in runs, and the attempts at more complex descriptions of sequential learning have been limited to simple run-run contingency assumptions (Butler, Myers, and Myers, 1969; Gambino and Myers, 1967; Myers, 1970; Rose and Vitz, 1966). Braine's general approach suggests greater complexity, since it assumes that people perceptually differentiate and generalize the temporal order of occurrence of units that may themselves be runs or even groups of runs. The context of a three-unit sequence, for example, can be defined by the arrangement of three different 'word' categories.

But Braine makes no claim of any mechanism that would permit the discrimination of *contrasting* word orders involving the same word categories. It is in this regard that rule-oriented theories of sequential learning seem most powerful. Subjects' ability to abstract certain kinds of transformational rules in deterministic sequences has been demonstrated by Restle and Brown (1970). Their theory suggests that a subject first generates rules to predict component parts of a pattern and then combines these subpatterns systematically. However, a strict appli-

cation of their normative theory would prevent the assessment of a subject's ability to selectively attend to certain higher-order rules and to ignore lower-order rules that define run properties. This is because the rules that determine event runs in Restle and Brown's analysis are usually correlated with the rules that describe the overall pattern. If a sequence begins with two A events, for example, a transformational description of the sequence would permit only runs of A that are multiples of two.

It seems reasonable to expect that people can detect and abstract higher-order relationships between events in situations that discourage the preliminary mastery of lower-order rules. Subjects undoubtedly possess a hierarchy of strategies that reflect their attending to various properties of a sequence. Expectancies in the form of simple transformational rules may result in dominant strategies, but clearly at some point a subject must revise his strategies if they are unsuccessful, revise them in light of the event schedule itself.

The present study attempted to demonstrate that subjects do attend to and abstract complex contingencies between events in probabilistic, nondeterministic schedules. To do so, the rules involving contrasting temporal arrangements of two events in a probabilistic four-choice task were examined. The particular events involved in the rules, as well as the run properties of these events, were themselves irrelevant to subsequent accuracy, but a subject had to attend to the *temporal order* of E_1 and E_3 events in order to anticipate correctly the onset of a third, different event (E_2 or E_4). In each of two contrasting orders ($E_1E_3E_2$ and $E_3E_1E_4$), the intermediate events occurred with a run length of one, so that the task also required a subject to inhibit a strong repetition-response tendency caused by the fact that at other times the E_1 or E_3 events occurred with longer run lengths. The four-choice sequence (see Method, also Figure 1, below) therefore not only provided a critical test of run theory but also permitted assessment of our subjects' ability to abstract higher-order contrasting relationships among the events themselves within single sequences.

Previous attempts to pit run cues against nonrun cues have met with limited success (Wolin, Weichel, Terebinski, and Hansferd, 1965; Halpern, Lantz, and Schwartz, 1969). The present study used (pre)training sequences in which the temporal order of two events was a reliable cue for prediction of a third event; interference with this cue by run cues was also minimized by correlating the run lengths with the events involved and by permitting the subjects to keep track of the order of events. It was hypothesized that these procedures would facilitate the

subjects' shift of attention to event arrangements when those arrangements became the sole basis for prediction in the test sequence.

METHOD

—Procedure—Three experimental groups received 192 training trials; a fourth received no training. All groups received 312 test trials. Both training and test situations were four-choice tasks in which each subject was instructed to predict which of four lights would occur on each trial. In a modified self-paced procedure, he was permitted up to 2 min after the onset of the signal light on each trial to register a prediction, although most responses were made within a few seconds. Other aspects of the trial cycle were controlled such that an event light of 2.1-sec duration occurred 1.4 sec after the subject's response. The experimenter manually controlled trial onset and programmed the appropriate event light for each trial according to a predetermined sequence.

—Apparatus—The subject's booth contained a box with a front panel set with four different colored lights, four corresponding response buttons, and a white center-mounted signal light. Control apparatus was located in an adjacent room. Hunter timers controlled the onset and duration of the event lights within a trial.

—Subjects—Sixty-four undergraduate female students from introductory psychology courses at Ohio State University served as subjects.

—Training conditions—The subjects were randomly assigned in equal numbers to one of four training conditions (P_1 , P_5 , R, and N), which differed according to characteristics of the training sequences and use of memory aids.

The two experimental groups were groups P_1 and P_5 . These subjects received patterned event sequences in which runs of lengths 1, 3, 5, and 7 were correlated with specific event lights so that whenever position E_i occurred, it was always correct for the same number of successive trials. Within the 192 trials, each run length appeared an equal number of times, as did each pair of run-length sequences; runs of length 1, say, were followed by runs of lengths 3, 5, and 7 an equal number of times. In both groups the onsets of some event runs were predictable from the relative order of *two* preceding event runs. The P_1 condition differed from that of P_5 in that the second event in the three-run contingency always occurred with run length of 1, while in the P_5 condition the intervening event had length run of 5; thus, subjects in groups P_1 and P_5 were required to use either the preceding two or six events to pick up the contingencies. The contingencies for groups P_1 and P_5 are shown in Table 1. In both conditions the joint occurrences of the specified event runs uniquely determined the occurrence of a third event run. The other two-run sequences were not followed by unique events but by a randomly determined event. Both groups were permitted to keep track of the response-event sequence with paper and pencil; no suggestions of coding techniques were given. Four different counterbalancing assignments of events (E_1 – E_4) to button and light positions were used for groups P_1 and P_5 , sub-

jects being assigned randomly in equal numbers to counterbalancing conditions.

Table 1. Training-condition contingencies for groups P_1 and P_5

Group	Contingency		
P_1	3, 1	→	5
	5, 1	→	7
	7, 1	→	3
P_5	1, 5	→	3
	3, 5	→	7
	7, 5	→	1

Note: Each contingency ($x, y \rightarrow z$) means that a run of length y preceded by a run of length x was always followed by a run of length z .

The two control groups were *groups N and R*. The subjects in group R received run lengths of 1, 3, 5, and 7, but the run lengths were randomly assigned to the four event-light positions, with the restrictions that each event light have an equal number of runs of a given length. No contingencies were introduced between lights or run lengths, and no memory aids were permitted these subjects. The subjects in group N received no training.

—Test sequence—All four groups received 312 trials on the same test sequence, with no memory aids permitted. Run lengths specific to the test sequence were 1, 2, 3, 4; only events E_2 and E_4 were assigned particular run lengths (2 and 4 respectively). Any of the four run lengths could occur at positions E_1 and E_3 . The structure of the test sequence is illustrated in Figure 1.

The following contingent rules were used (see Figure 1). Light E_1 with a randomly determined run length ($k = 1, 2, 3$, or 4) was always followed by E_3 with run length 1, and this joint occurrence was always followed by E_2 with a fixed run length of 2. Light E_3 with randomly determined run length (k) was always followed by E_1 with run length 1, and this joint occurrence was always followed by E_4 with a fixed run length of 4. After a run of E_2 or E_4 , either an E_1 or E_3 run of random length k could start.

In the test sequence, the onset of E_2 and E_4 were completely predictable using the above rules; the continuation and termination of those events were completely predictable from spatial or run-length cues. The onset and termination of E_1 and E_3 were sometimes predictable and sometimes not predictable. Again, four different counterbalancing orders of event (E_1 – E_4) assignments to button positions were used, with subjects randomly assigned in equal numbers to each counterbalancing condition.

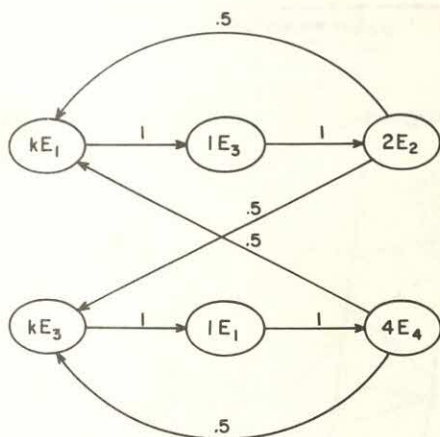


Fig. 1. Structure of the test sequence; prefixes in ovals show run length, arrows indicate events following each event run and the associated probabilities

RESULTS

Training trials

The main differences between groups R and P_1 and P_5 were, as expected, that the subjects using memory aids and correlated run-rule contingencies were superior in predicting the appropriate event runs. Figure 2 shows the asymptotic (last 50 trials) probability of an A_i response at onsets and within the several runs for the P_1 and P_5 groups. These subjects predicted both the continuation and termination of a specified run fairly well. Predictions of onsets ($k = 0$) for these groups were not so accurate, but were generally more accurate than corresponding onset predictions for subjects in group R. The latter group predicted onsets with well below chance accuracy (chance equals approximately .25).

Test trials

Results from the test trials involve predictions of the onset, continuation, and offset of the predictable runs. For clarity, data are collapsed across similar conditions and event types in the major analyses. For example, data from the two pattern-trained groups, P_1 and P_5 , were combined, since these groups did not differ in any important way on test trials. Similarly, responses to the predictable E_2 and E_4 events for all

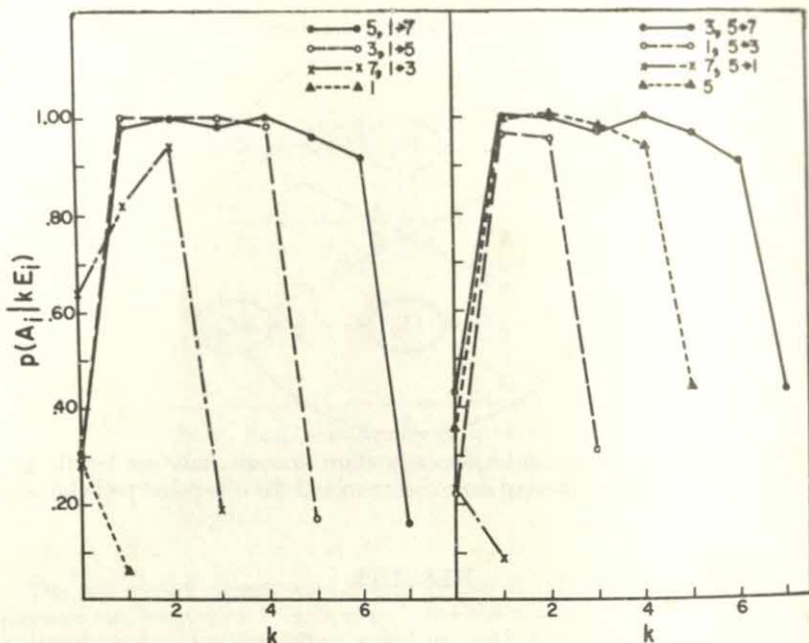


Fig. 2. Asymptotic probability of an A_i response within a run of E_i events in training; group P_1 in left panel, group P_5 in right panel. Performance on intervening runs ($\blacktriangle$ — — $\blacktriangle$) of lengths 1 and 5 for groups P_1 and P_5 respectively is also shown.

groups were pooled. Although subjects responded more accurately to all aspects of E_4 events, there were no interactions between this effect and any of the effects of primary interest.

Figure 3 shows the probability of various possible responses at the onset of the predictable E_2 and E_4 runs. The probability of a correct response at these points was around chance ($p = .25$) during the first block of approximately 50 trials and climbed steadily during the remaining trial blocks. The planned contrast between groups P_1 and P_5 and groups R and N was significant [$F(1, 60) = 5.29, p < .025$]; the pattern-trained groups predicted the onsets of E_2 and E_4 more accurately than did the other two groups. An examination of responding in the first trial block also reveals that group N was superior to the others, a finding which reflects initial negative transfer from the training sequences. Negative transfer was most apparent in group R , which showed a stronger tendency toward perseverative errors than did other groups.

All subjects were highly accurate in predicting the continuation of a

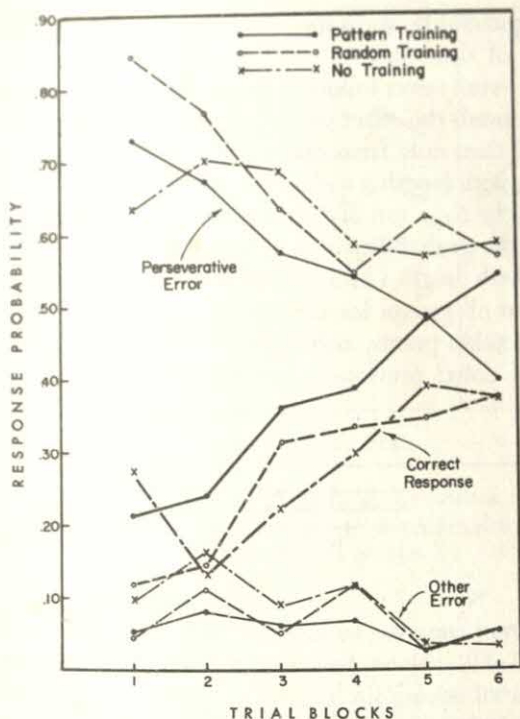


Fig. 3. Probability of various responses at start of predictable run (E_2 and E_4 combined); all groups

run, the probability of a correct response in general exceeding .80 asymptotically. The subjects in group N revealed some initial difficulty relative to the other groups. An analysis of variance of correct responses showed that the interaction of groups and trials was significant [$F(15, 300) = 8.99, p < .001$], as were the main effects of groups and of trials [$F(3, 60) = 4.39, p < .01$; and $F(5, 300) = 23.0, p < .001$, respectively]. Planned comparisons showed that groups P_1 and P_5 were significantly better than groups R and N [$F(1, 60) = 9.62, p < .01$], in their accuracy of predicting run continuations.

Because E_2 and E_4 events occurred only in runs of length 2 and 4 respectively, the *offset* of either of these runs was also predictable. Figure 4 shows the probability of predicting either of the appropriate events (E_1 or E_3) at the offset of E_2 and E_4 runs, as well as the probabilities of the two types of error responses. Clearly, the main difficulty at run terminations is reflected in the tendencies to persevere errone-

ously, which, predictably, were the greatest for group R. One very interesting aspect of these data is the fact that all groups learned very early that an E_4 event never followed an E_2 event and vice versa.

On the other hand, the offset of runs E_1 and E_3 was only sometimes predictable, and then only from event-arrangement cues—such as that after a run of E_1 with length k a run of E_3 was always of length 1, while after a run of E_2 or E_4 , a run of E_1 could be of any length from 1 to 4. Figure 5 shows the probability of a perseverative response following a run of E_1 or E_3 with length 1 at predictable and at unpredictable points. It also shows that all groups learned to discriminate between predictable and unpredictable points, although the probability of a perseverative response, as noted previously, remained fairly high at both. The

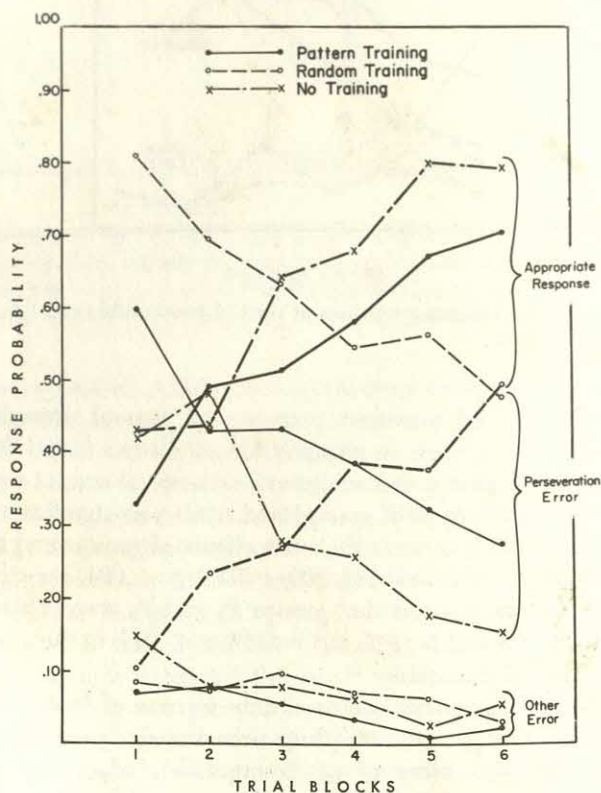


Fig. 4. Probability of various responses at end of predictable run (E_2 and E_4 combined); all groups

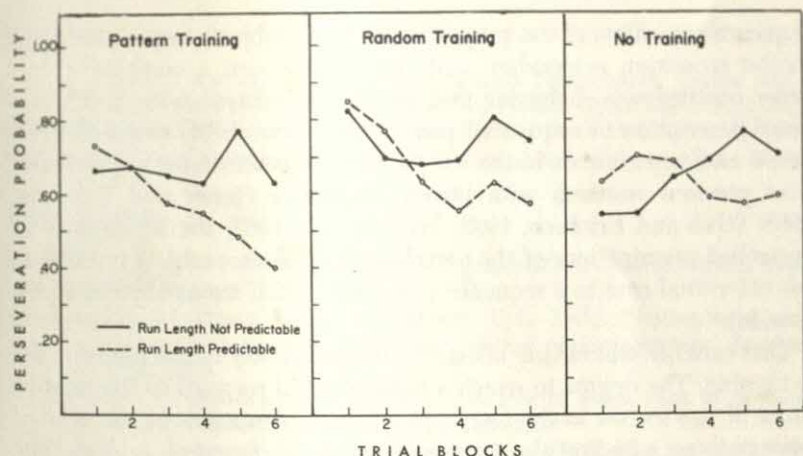


Fig. 5. Perseveration probability after E_1 or E_3 with a run length of 1 at points in the sequence where the run length is predictable (always length 1) or unpredictable (length 1, 2, 3, or 4); all groups

most important result of the analysis of variance here was the significant interaction of predictability and trials [$F(5, 300) = 17.12, p < .001$], which is quite clear in Figure 5. All groups initially perseverated more when the E_1 or E_3 run was predictable at length 1, but with additional training they learned to inhibit repetition responding at this point and respond more appropriately.

DISCUSSION

All groups learned to anticipate run onsets and to discriminate predictable from unpredictable run terminations on the basis of rules that used contrasting temporal-event orders in the test sequence. These findings are difficult for run theories that assume that the probability of a repetition response at onset points is a function only of the overall likelihood that the preceding event run continues (Myers, 1970; Restle, 1961, p. 235). Versions of sequential pattern learning that incorporate run-generalization and run-run contingencies (Butler, Myers, and Myers, 1969; Rose and Vitz, 1966) suffer the same limitations in explaining these findings, as does a contextual-generalization model.

None of those theories permit the subject to discriminate and abstract contrasting temporal arrangements of three events within the same

sequence regardless of run properties. But the subjects in this study did inhibit repetition responding and respond to such a complex event-order contingency—behavior that is more consistent with a rule-oriented description of sequential prediction. Their ability to predict run onsets and terminations in the test sequence demonstrates more clearly than previous research with binary sequences (Jones and Erickson, 1968; Strub and Erickson, 1968; Wolin et al., 1965) the importance of controlled examinations of the variables that affect a subject's ability to use contextual cues in a sequence (see Jones, 1971, for an expanded discussion).

One variable that clearly affects a subject's ability to use nonrun cues is training. The degree to which a subject could respond to the relative order of two events in the test sequence was a function of the type of prior training, a finding also consistent with a rule-learning analysis. The subjects in the two pattern-trained groups were significantly more effective in the test sequence than the control groups at eliminating perseverative errors and in learning to predict run onsets. The random-trained subjects were characterized by high perseverative responding, behavior more consistent with predictions of run theory. Interestingly, the superiority of groups P₁ and P₅ relative to groups R and N remained great over a major part of the test sequence, despite the fact that the two pattern groups did not show immediate positive transfer relative to group N. The overall superiority of the two pattern-trained groups is attributable to their effectiveness in overcoming perseverative tendencies.

Significant differences in performance between the pattern-trained groups and groups R and N show that subjects may be trained so that they are less likely to attend only to runs of homogeneous events. The picture that emerges from the performance of group R suggests that random structures encourage subjects to attend only to the ongoing run or immediately preceding event and that shifting this basis of attention is quite difficult.

The rules that the subjects discriminated in the present study demonstrate an ability to abstract from complex sequences the properties that define higher-order relationships between the events themselves. It is clear that most versions of run-theory cannot account for these subjects' performance in the test sequence, and indeed, a simple contextual-generalization approach is not sufficient. Rule-oriented descriptions such as that of Restle and Brown (1970) are most promising in this regard, although their approach needs to be expanded to account for attention processes that must come into play in sequences that contain both rule-determined and probabilistically determined events.

Notes

This research was carried out at the Human Performance Center and was supported in part by National Institutes of Health Grant MN12113. Reprint requests should be sent to Mari R. Jones, Ohio State University, 404C West 17th Avenue, Columbus, Ohio 43210. Received for publication June 22, 1971.

References

- Braine, M. D. S. 1963. On learning the grammatical order of words. *Psychological Review* 70:323-348.
- Butler, P. A., Myers, N. A., and Myers, J. L. 1969. Contingencies among event runs in binary predictions. *Journal of Experimental Psychology* 79:424-429.
- Gambino, B., and Myers, J. L. 1967. The role of event runs in probability learning. *Psychological Review* 74:410-419.
- Halpern, J., Lantz, A. E., and Schwartz, J. A. 1969. Cue properties of the event run in choice discrimination learning. *Journal of Experimental Psychology* 80:237-242.
- Jones, M. R. 1971. From probability learning to sequential processing: A critical review. *Psychological Bulletin* 76:153-185.
- Jones, M. R., and Erickson, J. R. 1968. Patterns of event runs in probability learning. *Psychonomic Science* 13:317-318.
- Myers, J. L. 1970. Sequential choice behavior. In *The psychology of learning and motivation: Advances in research and theory*, vol. 4, ed. G. H. Bower and J. T. Spence, pp. 109-170. New York: Academic Press.
- Restle, F. 1961. *The psychology of judgment and choice*. New York: Wiley.
- Restle, F. 1967. Grammatical analysis of the prediction of binary events. *Journal of Verbal Learning and Verbal Behavior* 6:17-25.
- Restle, F., and Brown, E. 1970. The organization of serial learning. In *The psychology of learning and motivation: Advances in research and theory*, vol. 4, ed. G. H. Bower and J. T. Spence, pp. 249-331. New York: Academic Press.
- Rose, R. M., and Vitz, P. C. 1966. The role of runs in probability learning. *Journal of Experimental Psychology* 72:751-760.
- Strub, M. H., and Erickson, J. R. 1968. Second-order dependency in probability learning. *Journal of Experimental Psychology* 78:261-268.
- Wolin, B. R., Weichel, R., Terebinski, S. J., and Hansferd, S. J. 1965. Performance on complexly patterned binary event sequences. *Psychological Monographs* 79(7, whole no. 600).

continued from page 248

- Whitlock, F. A. *Death on the Road: A Study in Social Violence*. New York: Barnes & Noble, 1971. Pp. xii, 212. \$9.25.
- Winter, Arthur, ed. *The Surgical Control of Behavior*. Springfield, Ill.: Charles C Thomas, 1971. \$8.00.
- Wisse, Ruth R. *The Schlemiel as Modern Hero*. Chicago: University of Chicago Press, 1971. Pp. xi, 134. \$5.45.
- Younghusband, Eileen; Birchall, Dorothy; Davie, Ronald; Pringle, M. L. Kellmer, eds. *Living with Handicap: The Report of a Working Party on Children with Special Needs*. Brooklyn: Gordon and Breach, 1971. Pp. xvii, 372. \$14.50.
- Zaetz, Jay L. *Organization of Sheltered Workshop Programs for the Mentally Retarded Adult*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xi, 229. \$12.75.

PROBABILITY PERFORMANCE: REINFORCEMENT PROCEDURE AND NUMBER OF ALTERNATIVES

Morton W. Weir

University of Illinois at Urbana-Champaign

With the number of response alternatives (two or four) and the type of informative feedback (contingent or noncontingent) manipulated in a probabilistic task, there were fewer choices of the high-probability alternative in the four-alternative contingent condition than in the two-alternative one, but the reverse was true in the two noncontingent conditions. The amount of information transmitted on each feedback trial may explain the results.

In two experiments using young children as subjects, Weir and Gruen (1965) reported that the frequency of choice of the high-probability alternative was lower in a three-alternative than in a two-alternative probabilistic task. This result is in direct opposition to those of a number of previous studies that have reported an increased frequency of choice of the high-probability alternative as the number of alternatives in the task was increased (Cotton and Rechtschaffen, 1958; Gardner, 1957, 1958, 1961). However, the Weir and Gruen result is in agreement with other unpublished findings of the present author, as well as those of Detambel (1955).

An examination of each of these experiments revealed that virtually all of them reporting more frequent choice of the high-probability alternative with increased numbers of alternatives made use of a *noncontingent* reinforcement procedure (more precisely, a noncontingent informative-feedback procedure). Thus, after each choice response, the subject was told which response was correct on that trial, regardless of which response he had made. That is, in the noncontingent procedure, the information the subject receives is independent of his response. Under a *contingent* informative feedback procedure, when the subject makes a correct response, he is told so; when he makes an incorrect response, he knows he was incorrect but is not told which response would have been correct on that trial. With such a procedure, it has been found that increased numbers of alternatives result in a decreasing proportion of choices of the high-probability alternative.

Several studies have actually included the contingent and noncon-

tingent procedures in a single design in which the number of alternatives was also a variable. But the results have been inconsistent and inconclusive. Detambel (1955), using tasks with two, four, or eight alternatives, found that the contingent procedure produced a decreasing frequency of choice of the high-probability alternative as the number of alternatives increased but found no effect of number of alternatives in the noncontingent case. Both of these results were predicted by the mathematical model that Detambel outlines in that paper.

On the other hand, Neimark (1956), using two- and three-alternative tasks, reported no difference between contingent and noncontingent procedures, while Gardner (1957) reported that in both the contingent and the noncontingent cases, use of a three-alternative task resulted in more choices of the high-probability alternative than did use of a two-alternative task. In attempting to explain the discrepancy between his results and those of Detambel and Neimark, Gardner states that the other two investigators had run too few trials to obtain the effect, but he ignores in his explanation the fact that Detambel had obtained results *opposite* to his own when both used a contingent procedure.

In addition, Gardner's findings are subject to question because the two-alternative data that he used to compare with data from the three-alternative contingent task had been obtained in an earlier *noncontingent* experiment. He explains this choice of comparison groups by saying that in the two-alternative task there should be no difference between the contingent and noncontingent procedures. Although this is a reasonable assumption, it is not necessarily correct. In any event, comparing the results of one group of subjects run in an earlier experiment with another group from a later one is in itself somewhat risky.

The current status of the combined role of number of alternatives and contingent versus noncontingent informative-feedback procedures is therefore unclear, to say the least. This lack of clarity, combined with four experiments by the present author which show that increasing the number of alternatives in a contingent task produces a decrease in the frequency of choice of the high-probability alternative (a finding reported previously only by Detambel, 1955, and largely ignored by later researchers), prompted the present experiment. Since our previous data came almost entirely from children, and those of the other studies almost entirely from adults, it was decided to manipulate the number of response alternatives (two versus four) and type of reinforcement procedure (contingent versus noncontingent) in a single design using only adult subjects.

METHOD

—Subjects—The subjects were 94 students in the introductory psychology class at the University of Illinois. Each was selected on the basis of having a time period free when the experimenter did. There were 69 females and 25 males in the sample, and assignment to experimental groups was random. Inspection of the data after the experiment revealed no consistent sex effects among the four experimental conditions, although the males on the whole tended to choose the high-probability alternative somewhat more frequently than did the females.

—Apparatus and procedure—The apparatus consisted of a yellow panel 18 in. high and 30 in. wide, in which was centered a horizontal row of four black buttons $2\frac{1}{2}$ in. in diameter and mounted 3 in. apart. Three inches above each button was a green jeweled light 1 in. in diameter, and centered above these four lights was a red jeweled signal light, 2 in. from the top of the panel. The apparatus was constructed so that a yellow wooden shield could be placed over the outside two buttons and lights, converting it from a four- to a two-alternative apparatus.

Subjects arrived individually and were assigned to one of the four experimental conditions. All were instructed that this was a task in which they were to predict which one of the two (or four) green lights would come on, and that they could make their prediction by pressing the button below the light which they expected would come on. They were informed that the red signal light would come on when it was time to make a choice. The subjects in the noncontingent condition were then told: "On each trial, you will know which button was correct by seeing which light actually went on." The subjects in the contingent condition were told: "If you predict correctly, the light above the button you pressed will come on; if your prediction is incorrect, no light will come on."

Half of the subjects in the contingent and noncontingent groups were given a two-alternative task; the other half had the four-alternative task. In the two-alternative task, one of the lights was programmed to occur on 64% of the trials, and the other on 36% of the trials. In the four-alternative task, the proportions were 64%, 12%, 12%, and 12%. In the contingent condition the subjects saw a green light illuminated only when they actually made the correct response for that trial. Thus, in the contingent condition, an incorrect response resulted in a blank trial in the sense that no information light occurred. The green information lights were programmed to occur (when appropriate) immediately after the choice response and to remain on for 2 sec. Termination of the green light was followed by a 4-sec intertrial interval, after which the red signal light occurred and remained on until the next choice response was made.

The reinforcement schedule was random within the above percentage restrictions, and each subject received the same random schedule (that is, for all subjects, the 64% information light was scheduled to occur on the same trials). To generate the reinforcement schedule for the three less frequently occurring lights in the four-alternative task, the 36% occurrence of the least

frequent light in the two-alternative task was split equally among these three alternatives. Approximately equal numbers of subjects in the two-alternative task had the high-probability alternative on the right and left; in the four-alternative task, the 64% alternative occurred in each of the four positions about equally among subjects. Fifty subjects served in the two-alternative condition, half in the contingent and half in the noncontingent groups. Forty-four subjects served in the four-alternative task, half in each of the contingency conditions. Each subject completed 200 trials.

RESULTS AND DISCUSSION

A 2 (two versus four alternatives) $\times$ 2 (contingent versus noncontingent) $\times$ 10 (blocks of 20 trials) analysis of variance was performed on the number of choices of the high-probability alternative. The analysis revealed a significant main effect associated with trials [$F(9, 810) = 15.03, p < .001$] and a significant interaction between trials and number of response alternatives [$F(9, 810) = 6.61, p < .001$]. In addition, there was also a significant interaction between contingency condition and number of response alternatives [$F(1, 90) = 10.26, p < .001$]. All three of these significant effects may be seen in Figure 1, which presents the frequency of choice of the high-probability alternative in blocks of 20 trials for the four experimental groups.

The interaction between blocks of trials and number of response alternatives was the result of a generally increasing trend in number of choices of the high-probability alternative in the four-alternative con-

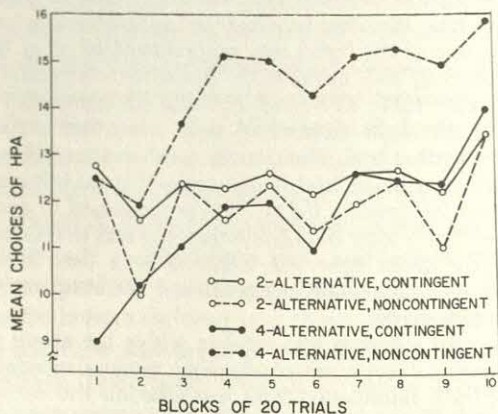


Fig. 1. Mean number of choices of the high-probability alternative (HPA) per block of 20 trials for the four experimental groups

ditions, with little or no increase in the number of those choices as trials progressed in the two-alternative conditions.

Of greater importance to the original purpose of this study, however, is the interaction between contingency conditions and number of response alternatives. In the noncontingent groups (broken lines, Figure 1), the frequency of choice of the high-probability alternative was considerably higher in the four- than in the two-alternative condition. This is the finding reported frequently by Gardner (1957, 1958, 1961) for the noncontingent case. In the contingent groups, however (solid lines, Figure 1), the frequency of choice of the high-probability alternative was lower for the four- than for the two-alternative conditions, a finding in agreement with the prior findings of Detambel (1955) and Weir and Gruen (1965) for the contingent case. It should be noted that frequency of choice of the high-probability alternative was lower for the four-alternative contingent group than for the two-alternative contingent group only through approximately the first half of the task. In the first of the Weir and Gruen experiments, a similar finding occurred over the course of 192 trials for two- and three-alternative tasks.

A Tukey (a) test was performed on the total number of choices of the high-probability alternative by the four main experimental groups, in order to examine in more detail the source of the interaction between contingency condition and number of response alternatives. This analysis revealed that the four-alternative noncontingent group was significantly higher in its choices of the high-probability alternative than any of the other groups, while the four-alternative contingent group was significantly below the two-alternative contingent group (all $ps < .01$). Thus, the effect of increasing the number of response alternatives in the contingent case was opposite to that in the noncontingent case. Choices of the high-probability alternative increased with increasing numbers of alternatives in the noncontingent case but decreased with increasing number of alternatives in the contingent situation. And, as Gardner (1957) had assumed, there was no effect of the contingency variable in the two-alternative task, at least as far as choices of the high-probability alternative were concerned.

Although the pattern of these results is clear, the explanation for the finding that varying the number of response alternatives has an opposite effect in the contingent and noncontingent situations is not so clear. Taking the noncontingent case first, a higher frequency of choice of the high-probability alternative in the four-alternative as compared with the two-alternative task may have been due to the greater contrast in the percentage payoff of the alternatives in the four-alternative task. In the

four-alternative case, there is a difference of 52% in the frequency of occurrence of the high-probability alternative (64%) compared with any of the low-probability alternatives, each of which occurs 12% of the time. In the two-alternative task, this discrepancy is only 28% (64% versus 36%). Other research (Weir, 1964, 1965) has shown that adults are sensitive to differences as small as 10% in the payoff schedule, and if the informative feedback is noncontingent, they should be aware of this discrepancy in frequency of occurrence quite early in the task.

In the contingent case, on the other hand, a considerable number of trials are included on which the subject receives no information as to which alternative is actually correct. In the two-alternative case, this should make relatively little difference (see Figure 1), since if the choice of one alternative is incorrect, the other will probably be assumed to have been correct. But when more than two alternatives are involved, the noninformative nature of an incorrect trial could be expected to play a significant role. With four alternatives from which to choose, a high proportion of the choices at the beginning of the task will be to one of the three low-probability alternatives. And, 88% of the time, such a choice will provide no information as to which alternative *was* correct on those trials.

If subjects are sensitive to runs of events, and if these enter into their hypotheses about the 'solution' of a probabilistic task (Stevenson and Weir, 1963), then trials on which no event occurs should slow the process of testing such hypotheses. These noninformative trials are equivalent to the 'blank' trials (trials on which no alternative is scheduled to pay off) described by Neimark (1956), who found their effect to be one of depressing the slope of the learning curve. And since any contingent procedure with more than two alternatives results in blank trials from the subject's point of view, then such a procedure should depress the number of choices of the high-probability alternative, as it obviously does (see groups four-alternative contingent versus four-alternative non-contingent, Figure 1).

The explanation of the pattern of results in this study may therefore be summarized as follows. Ordinarily, choice of the high-probability alternative will increase whenever the contrast in percentage payoff between it and the other alternatives increases. Thus, when performance in a two-alternative task with a 64:36% reinforcement distribution is compared with that in a four-alternative task with a 64:12:12:12% distribution, the latter should show more frequent choice of the high-probability alternative. However, adding noninformative (blank) trials to the four-alternative task, as was true for the contingent condition in this

study, masks this percentage contrast, particularly in the early trials. This masking effect of the blank trials affects performance in the four-alternative situation but not in the two-alternative, as subjects in the latter task can simply assume that the other alternative would have been correct on any blank trial, and thereby lose no information.

Such an explanation would predict that choice of the high-probability alternative in the four-alternative contingent condition should be quite low to begin with, when compared with the other three groups, but should improve steadily as the subjects get information about the contrast in percentage payoff. As this information accumulates, choice of the high-probability alternative in the four-alternative contingent condition should surpass that in the two-alternative conditions, as the contrast in percentage payoff is greater in the four- than in the two-alternative task. This prediction is confirmed by examination of Figure 1. It is also likely that the difference in level of choice of the high-probability alternative between the four-alternative contingent and noncontingent groups would have disappeared if the task had been continued over a longer series of trials.

Another test of the proposed explanation would involve a comparison of the level of responding with the high-probability alternative in the four-alternative contingent and noncontingent conditions as a function of the number of prior informative-feedback trials (an informative-feedback trial being defined as a trial on which a green information light occurs). That is, if the subjects in the four-alternative contingent condition respond less with the high-probability alternative (see Figure 1) because they have less informative feedback than subjects in the four-alternative noncontingent group, then the performance of the two groups should be quite similar if the amount of informative feedback could somehow be equated.

Now, in the four-alternative noncontingent condition, since informative feedback occurs on each trial, the level of responding with the high-probability alternative at the point of a subject's n th feedback trial may be ascertained by examining the level of that responding on that subject's n th trial in the task. However, the n th feedback trial for a subject in the four-alternative contingent condition will ordinarily occur much later than the n th task trial. The question then is, Will the level of responding with the high-probability alternative by these two subjects be the same at the time of the n th feedback trial, regardless of the point in the task at which that trial occurs?¹

In order to answer that question, feedback trials 10, 20, 30, and so on, were located for each subject in the four-alternative contingent condi-

tion, and his levels of responding with the high-probability alternative at those points were determined by counting those responses during the ten trials surrounding those points. For example, if his feedback trial 20 occurred at task trial 51, his number of responses with the high-probability alternative on task trials 47 through 56 was recorded. The same procedure was used for each subject in the four-alternative noncontingent condition, with the exception that the location of the n th feedback trial was always at task trial n .

The results of this procedure are shown in Figure 2. When the data are examined in this manner, there was very little difference in the level of responding with the high-probability alternative between subjects in the four-alternative contingent versus noncontingent conditions. The numbers in parentheses in Figure 2 indicate the number of subjects represented at each point in the contingent group. (All 22 subjects in the noncontingent condition are represented at all points along that group's curve). The graph does not continue beyond feedback trial 90, as only two subjects in this contingent group had 100 or more instances of informative feedback in the 200 task trials.

An analysis of variance was performed on a subset of the data shown in Figure 2. The number of choices of the high-probability alternative

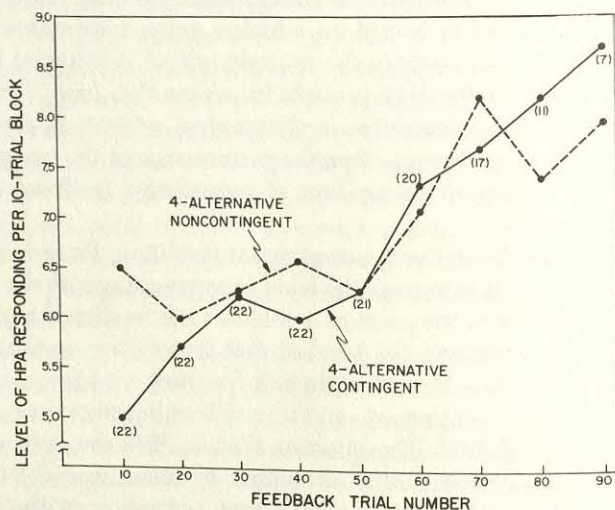


Fig. 2. Mean number of choices of the high-probability alternative (HPA) in the 10 trials surrounding every tenth feedback trial; numbers in parentheses are number of subjects at each point in the four-alternative contingent condition

by the 20 subjects in the four-alternative contingent condition who had at least 60 trials ending in informative feedback was included in the analysis, along with the performance of the 22 subjects in the four-alternative noncontingent group. This analysis indicated no difference as a function of contingency condition ($F < 1$) and no interaction between contingency condition and feedback trial number [$F(5, 200) = 2.16, p > .05$]. The main effect associated with trials was significant [$F(5, 200) = 6.32, p < .01$]. Thus, both Figure 2 and this analysis confirm the hypothesis that the major factor underlying the difference between the level of responding with high-probability alternative in the four-alternative contingent and noncontingent groups (shown graphically in Figure 1) is that information on the relative payoff percentages accumulates slowly in the contingent group, presumably because of the high frequency of blank trials. Interestingly, Figure 2 indicates that adults in the four-alternative contingent group were apparently able to accumulate this information as accurately as those in the four-alternative noncontingent condition, even though frequent blank trials intervened between instances of informative feedback.

Notes

This investigation was supported by PHS Research Grant HD00874 from the National Institute of Child Health and Human Development. I thank Judith Cohen Conger for her aid in data collection and analysis. Received for publication June 14, 1971.

1. I am indebted to my colleague, Dr. Victoria Seitz, for suggesting this method of examining the data.

References

- Cotton, J. W., and Rechtschaffen, A. 1958. Replication report: Two- and three-choice verbal conditioning phenomena. *Journal of Experimental Psychology* 56:96.
- Detambel, M. H. 1955. A test of a model for multiple-choice behavior. *Journal of Experimental Psychology* 49:97-105.
- Gardner, R. A. 1957. Probability learning with two and three choices. *American Journal of Psychology* 70:174-185.
- Gardner, R. A. 1958. Multiple-choice decision behavior. *American Journal of Psychology* 71:710-717.
- Gardner, R. A. 1961. Multiple-choice decision behavior with dummy choices. *American Journal of Psychology* 74:205-214.
- Neimark, E. D. 1956. Effects of type of nonreinforcement and number of alternatives in two verbal conditioning situations. *Journal of Experimental Psychology* 52:209-220.

- Stevenson, H. W., and Weir, M. W. 1963. The role of age and verbalization in probability-learning. *American Journal of Psychology* 76:299-305.
- Weir, M. W. 1964. The effect of patterned partial reinforcement on children's performance in a two-choice task. *Child Development* 35:257-264.
- Weir, M. W. 1965. Children's behavior in a two-choice task as a function of patterned reinforcement following forced-choice trials. *Journal of Experimental Child Psychology* 2:85-91.
- Weir, M. W., and Gruen, G. E. 1965. Role of incentive level, number of choices, and type of task in children's probability learning. *Journal of Experimental Child Psychology* 2:121-134.

THE EFFECT OF HYPNOTIC AGE REGRESSION ON SIZE CONSTANCY

Herschel W. Leibowitz, Charles Graham, and Michael Parrish

Pennsylvania State University

Hypnotic age regression of college-age adults to ages nine and five produced a lowering of size constancy in a manner typical of children those ages. But the effect was not significantly different from that produced by nonhypnotized controls instructed to imagine they were nine and five. The multiplicity of mechanisms that may account for size constancy is emphasized in interpreting these data: though an individual cue such as linear perspective may be affected by age-regression techniques, the size-constancy effect, apparently subserved by a number of mechanisms, is not susceptible.

It has recently been demonstrated that the magnitude of both the Ponzo and Poggendorff illusions is altered by hypnotically induced age regression. During regression to age nine, the responses of college-age observers to both illusions were equivalent to those made by nine-year-old children. With further regression to five years of age, the illusion magnitudes were altered even more, approaching but not reaching the normative values typical of five-year-olds (Parrish, Lundy, and Leibowitz, 1968, 1969).

It has also been suggested that both size constancy and the Ponzo illusion are mediated by a similar mechanism (Fisher, 1968). Specifically, the individual learns through experience in the normal environment that converging lines are associated with distance. In a normal viewing situation, this process provides a cue to distance and permits the necessary compensation for the diminished retinal-image size of distant objects, thereby producing size constancy. However, with the two-dimensional Ponzo figure, the converging lines falsely indicate distance and the correction is misapplied, producing an illusory effect.

The present experiment was concerned with whether the size-matching behavior of hypnotically age-regressed adults would, like their Ponzo-illusion behavior, be typical of matches made by younger children. Since both the Ponzo illusion (Leibowitz and Judisch, 1967) and size constancy (Leibowitz, Pollard, and Dickson, 1967) increase with chronological age, a determination that both measures can be reversed

by hypnotic age regression would be relevant to the assumption that the mechanisms underlying these two phenomena are similar.

Since size matches are sensitive to particular experimental conditions, a control group of children was tested under the same stimulus conditions as the college-age adults. Additionally, a nonhypnotized adult group was instructed to perform the task as they imagined children of the ages in question would, in order to determine whether the observed effects would also be obtained in the absence of hypnosis.

METHOD

—Subjects—The ten experimental observers (mean age, 19.0 years) and seven adult controls (mean age, 19.1 years) were students from elementary psychology classes who responded affirmatively to the statement "I would be willing to participate in a hypnosis experiment" and subsequently achieved the maximum objective score of 8 on the Barber Suggestibility Scale (Barber and Glass, 1962) during a preliminary screening procedure. The control group of children consisted of five five-year-olds and five nine-year-olds from the local community. All observers had normal vision with or without correction and were unaware of the purpose of the experiment.

—Apparatus—The standard stimuli were wooden circles painted flat white; they were 3.0, 12.0, and 20.0 in. in diameter, were viewed at observation distances of 22, 88, or 146.7 ft, and subtended a constant visual angle of .65 deg. The supports for the stimuli were constructed in proportion to the circle diameters, so that all aspects of the stimuli subtended the same visual angle at all observation distances (for details, see Leibowitz and Dato, 1966). The stimuli were shown, one at a time, to each observer, who was seated at one end of a 157-by-8-ft university corridor. The luminance in the corridor was 2 ftc. Ninety deg to the observer's left, at a distance of 22 ft, was the 4-by-8-ft panel, covered with black flock paper, that served as a background for the comparison stimuli. These latter were 36 circles cut from white cardboard and ranging in diameter from .5 to 6.0 in. in .5-in. steps, and from 6 in. to 30 in. in 1-in. steps. The comparison stimuli were fixed on the center of the background by means of velour tape.

—Procedure—The observer's task was to indicate whether the comparison object was bigger or smaller than the test object. The order of presentation of the comparison series was chosen so that the initial responses would vary from larger to smaller several times (to ensure that the instructions were understood). The remainder of the series was presented in random order, with the matched sizes being determined by interpolation between the points where the observer's responses changed.

The sequence of experimental conditions and the presentation order of the standard stimuli were randomly determined for each observer. The experimental group was tested weekly under one of four conditions: normal, hypnotized without age regression, hypnotic regression to age nine, and hypnotic

regression to age five. The control group of adults (the 'imagination,' or 'simulation,' controls) was tested under the same conditions, except that they were not hypnotized but were asked to perform the task as if they were nine- and five-year-olds. The control group of children was tested under the normal condition.

RESULTS

The mean matched-size values for all conditions of the experiment are summarized in Table 1. It will be noted, in line with previous observations, that the matches of children and adults were similar at the shorter observation distances. However, at the longest observation distance, the children showed less tendency toward size constancy than did the adults, with the discrepancy being greater for the younger children. As expected, the difference between children and adults in size constancy was manifest only at the longest observation distance, and it is here that the effect of any of the experimental manipulations should be apparent in the adult groups.

The performance of the adults in the experimental group paralleled that of the children. Adults who were hypnotically age-regressed made progressively smaller matches than they did under the normal or hypnotized-without-age-regression conditions. The imagination controls also produced size matches smaller than under normal conditions when asked

Table 1. Mean matched size in inches as a function of distance

	Test-object distance in feet (and diameter in inches)		
	22 (3)	88 (12)	146.7 (20)
Experimental group ($N = 10$)			
Normal	2.68	11.70	19.80
Hypnotized, no age regression	2.78	11.48	18.75
Regressed to age nine	2.60	11.28	17.70
Regressed to age five	2.73	10.93	15.50
Child controls			
Age nine ($N = 5$)	2.74	11.55	16.34
Age five ($N = 5$)	2.83	10.60	14.90
Adult controls ($N = 7$)			
Normal	2.69	12.62	21.17
Simulate nine years	2.79	11.96	18.71
Simulate five years	2.65	11.79	18.67

to simulate the behavior of children, but the decrease was essentially the same for both simulated age-levels.

Significant differences ($p < .05$) were found between regression to age five and simulation of age five ($t = 2.15$), as well as between simulation of age five and the five-year-olds themselves ($t = 2.43$). On the other hand, comparison of the matches between the adults hypnotically regressed to age five and the five-year-olds was not significant. Similarly, there were no significant differences between groups shown by the t tests for the nine-year-old condition.

However, since the adults in both the experimental and control groups departed from constancy with the manipulation of hypnotically regressed or imagined age and observation distance, the crucial comparison is the difference between the normal and regressed-to-age-five conditions for the experimental group as opposed to the difference between the normal and imagined-age-five conditions for the imagination controls. The critical test in this instance is evidence of a significant interaction between groups and regressed or imagined age-level. Analysis of variance based on these change scores yielded neither significant group differences ($p > .75$) nor significant interaction effects ($p > .75$). Thus, although hypnotic age regression resulted in lower matches at the greatest observation distance in a direction typical of young children, this departure from constancy cannot be attributed to the effect of the hypnotic suggestion.

DISCUSSION

The basis for the present experiment was the theoretical assumption that the Ponzo illusion and size constancy are subserved by the same distance-correction mechanism. Evidence has recently been obtained (Leibowitz, Brislin, Perlmutter, and Hennessy, 1969), consistent with the results of the present experiment, that this hypothesis may not take full account of the differences inherent in the two situations. The size-constancy effect, which by definition can occur only in three-dimensional space, is a more complex, multidetermined phenomenon, of greater magnitude than the two-dimensional Ponzo illusion. This is apparent in the normal viewing situation, where a size correction of orders of magnitude is needed to achieve size constancy, whereas in viewing the Ponzo illusion, a correction of the order of approximately 10% to 15% is typically observed.

Thus, a more parsimonious hypothesis is to assume that the cue of linear perspective, which is the basis for the Ponzo illusion, may be,

under certain circumstances, just one of the cues an individual uses in maintaining size constancy. The demonstration that the same hypnotic procedure which had produced significant age-regression effects for the Ponzo illusion (Parrish et al., 1968, 1969) did *not* do so for size constancy is consistent with this view. It is evident that the additional determinants of the size-constancy effect were not susceptible to the present hypnotic techniques. However, it is equally apparent that the two sets of investigations considered together indicate that (for reasons yet to be explained) at least one of the determinants, linear perspective, under certain conditions,¹ is subject to manipulation through hypnotic age-regression techniques.

Notes

The research was supported by Grant MH08061 from the National Institute of Mental Health. The writing of this report was supported in part by Contract Nonr 4731(00) from the Office of Naval Research to the Unit for Experimental Psychiatry, the Institute of the Pennsylvania Hospital. Address reprint requests to the first author, at the Department of Psychology, Pennsylvania State University, University Park, Pennsylvania 16802. Dr. Graham is now at the Unit for Experimental Psychiatry, Institute of the Pennsylvania Hospital and University of Pennsylvania; Dr. Parrish is now also at the University of Pennsylvania. Received for publication July 6, 1971.

1. For an alternative interpretation of the effect of age regression on the Ponzo and Poggendorff illusions (Parrish et al., 1968, 1969), see Ascher, Barber, and Spanos (1972).

References

- Ascher, L. M., Barber, T. X., and Spanos, N. P. 1972. Two attempts to replicate the Parrish-Lundy-Leibowitz experiment on hypnotic age regression. *American Journal of Clinical Hypnosis* 14:178-185.
- Barber, T. X., and Glass, C. B. 1962. Significant factors in hypnotic behavior. *Journal of Abnormal and Social Psychology* 64:222-228.
- Fisher, G. H. 1968. Illusions and size-constancy. *American Journal of Psychology* 81:2-20.
- Leibowitz, H. W., and Dato, R. A. 1966. Visual size constancy as a function of distance for temporarily and permanently monocular observers. *American Journal of Psychology* 79:279-284.
- Leibowitz, H. W., and Judisch, J. M. 1967. The relation between age and the magnitude of the Ponzo illusion. *American Journal of Psychology* 80:105-109.
- Leibowitz, H. W., Brislin, R., Perlmutter, L., and Hennessy, R. T. 1969. Ponzo perspective illusion as a manifestation of space perception. *Science* 166:1174-1176.

- Leibowitz, H. W., Pollard, S. W., and Dickson, D. 1967. Monocular and binocular size-matching as a function of distance at various age-levels. *American Journal of Psychology* 80:263-268.
- Parrish, M., Lundy, R. M., and Leibowitz, H. W. 1968. Hypnotic age-regression and the magnitude of the Ponzo and Poggendorff illusions. *Science* 159:1375-1376.
- Parrish, M., Lundy, R. M., and Leibowitz, H. W. 1969. Effect of hypnotic age regression on the magnitude of the Ponzo and Poggendorff illusions. *Journal of Abnormal Psychology* 74:693-698.

VERBAL BEHAVIOR AS A FUNCTION OF AMOUNT OF SCHOOLING

James W. Hall
Northwestern University

Of two groups of Colombian children differing by two years in amount of schooling but similar in age (mean = 10.4 years), the group with more schooling was markedly superior in paired-associate learning and slightly superior in verbal-discrimination learning. The groups did not differ in rate of forgetting of paired associates over one week. Only those with more schooling evidenced the use of associative and acoustic attributes in memory encoding. Word-association differences paralleled those typically found with age differences.

The primary objective of the research reported in this paper was to examine the relationship between early formal school experiences and the development of certain learning and memory processes. This was accomplished by comparing the performance of groups of children similar in age, but substantially different in amount of formal schooling, on tasks of recognition memory, paired-associate learning and retention, verbal-discrimination learning, and word associations.

The *recognition-memory task* was designed to test an 'encoding shift' hypothesis suggested by Underwood (1969) in a recent theoretical paper. In that article Underwood proposed that memory for verbal material be conceptualized as consisting of a number of attributes of the material that are recorded and stored in some fashion. Among the several attributes discussed were verbal associations and acoustic characteristics. It seems clear that mature human memory relies heavily on the verbal-associative attribute when meaningful verbal material is involved. Various studies (see Hall and Ware, 1968, for example) have demonstrated the importance of the associative attribute in the memory of children as well. In contrast, the acoustic attribute appears to be more important where less meaningful verbal material is involved, particularly in the case of short-term memory (Underwood, 1969). This reasoning led to a specific hypothesis of developmental changes in encoding for memory, namely, that during early childhood a shift occurs, from relatively little use of the associative attribute and greater use of the acoustic attribute, and toward greater importance of the associative attribute in relation to

the acoustic one. A recent experiment by Bach and Underwood (1970) appeared to confirm this hypothesis, showing the predicted shift in seven- to eleven-year-olds.

If verbal-associative factors do increase in importance during early childhood, along with a relative decline in the importance of sound patterns per se, formal educational experiences (particularly early school instruction in language and reading) might be important determinants of this shift. The recognition-memory task was designed as the most direct test of this encoding-shift hypothesis. It involved the presentation of a series of words, followed by a recognition test. In the recognition list were words that were highly similar acoustically to certain of the to-be-remembered words, and words that were highly related associatively to these same words. The relative frequency with which these two classes of words were falsely recognized (judged 'old') in relation to control words was considered to be a reflection of the importance of the two attributes in the memory process. Procedures similar to this have been used by several investigators to examine encoding processes (Underwood and Freund, 1969; Bach and Underwood, 1970; Hall and Crown, 1970).

The *paired-associate task* was used to examine possible differences between the two groups both in learning and in retention. It was anticipated that faster paired-associate learning would be found for the children with more schooling than for those with less. On the other hand, the current 'interference' theory of forgetting led to quite a different prediction about retention. To the extent that the children had a richer learning history, they also would have more sources of proactive interference. Thus, one might predict that if the two groups were brought to an equal level of learning, forgetting (due to proactive interference) would be greater for the children with more schooling. In an effort to equalize degree of initial learning, the two groups were given differential numbers of trials based on the results of pilot research. The result was an apparently highly similar level of learning for the two groups. One week later, subjects were given one test trial and one relearning trial on these same items to examine retention differences between groups.

The *verbal-discrimination task* was used because it appeared likely that the basis for effective performance on this task is quite different from that for high performance on the other tasks used. If it is the frequency attribute rather than acoustic or associative ones that is vital in verbal-discrimination learning, formal education might not have any substantial effect on this performance. For that reason, the verbal-discrimination task was included in the investigation.

Data from a *word-association test* were collected for two reasons. First, comparisons between the two groups in terms of certain response characteristics are interesting in themselves. It was anticipated, for example, that the subjects with more schooling would give a greater number of contrast responses. Greater response homogeneity also was expected for this group, and greater variability anticipated for the group with less schooling.

The second purpose of the word-association test was to provide data of use in the recognition task described earlier. In order that the associates employed there be equally strong for the two groups insofar as probability of occurrence in word-association testing was concerned, it was necessary to obtain word-association data before administering the recognition task.

METHOD

—Subjects—The subjects were drawn from a low socioeconomic area of Cali, Colombia. Information on approximately 80% of the families residing in this area had been obtained by the research staff of a nutrition project there. This information was used to identify some 350 children from 8 years and 5 months to 12 years of age. Next, through school and home visits, the educational status and background of these children were determined. Such data eventually were obtained for approximately 250 children. From this subject pool, pairs of children were matched such that the members of each pair were the same age (within two months) but one member had been in school at least two years longer than the other and was at least one grade higher. In this fashion two groups of 35 children each were formed, the one labeled the *high-schooling group* and the other the *low-schooling group*. At the time these children were identified, all were enrolled in school, so that the differences in amount of schooling were a function of differences in the ages at which they began school. Unfortunately, it was not possible to determine the reasons for the wide variation in starting ages. The implications of this unanswered question will be considered later.

—Tasks—With few exceptions, all the testing described below was done at a research station near the homes of the children. Most of the testing was done by Colombians who were members of the nutrition project's testing staff. All verbal materials were, of course, in Spanish, and were presented orally by the examiners. Extensive pretesting of materials and training of examiners was done before administration of the tests, which took three sessions. In the first session, word-association data were collected. The second session was several weeks later, and consisted of the verbal-discrimination task followed by paired-associate learning. The third session, which was exactly one week after the second, began with a test of retention of the paired-associate items, followed by the recognition-memory task.

In the *recognition-memory task*, four lists of words were used. Lists 1a and 2a each consisted of 12 common to-be-remembered words, and lists 1b and

2b each included 16 recognition-test items. The order of presentation was list 1a, followed, after a 1-min filler task, by list 1b as recognition test. Then list 2a was presented, followed in 1 min by list 2b. The recognition tests each included four categories of four words each. One of these categories consisted of words associatively related to four of the to-be-remembered words, and a second category consisted of words acoustically similar to the same four words. For example, in the case of the to-be-remembered word *HOMBRE* (MAN), the recognition test included the words *MUJER* (WOMAN) and *NOMBRE* (NAME). Also, on each recognition list were four 'old' words, that is, four of the actual to-be-remembered words. Finally, each recognition list contained four words that were unrelated to previously presented words. These unrelated words functioned as control items in that they provided a baseline of false-recognition frequency.

In total, then, the recognition lists contained eight 'old' words, eight associates, eight acoustically similar words, and eight unrelated words. The particular associates used were chosen on the basis of the word-association data. In each case, the associate selected was that word which had been given most frequently by both groups of children when its corresponding to-be-remembered word was presented as a stimulus item in the word-association test.

In the *paired-associate task*, four pairs of words were presented for learning. The stimulus items were common male names, and the response items were the four nouns *NARIZ*, *LIBRO*, *ÁRBOL*, and *TAZA* (NOSE, BOOK, TREE, and CUP). After an initial familiarization trial, the anticipation method of presentations was used, with an anticipation interval of 5 sec. The subjects in the high-schooling group were given three anticipation trials; those in the low-schooling group were given six such trials. This was done in an attempt to equalize the groups' initial learning so that their retention differences could be examined. The particular numbers of trials used were based on pilot data gathered earlier for a portion of the subjects in each group.

Exactly one week after paired-associate learning, each subject was brought in for a *retention test*. One test trial was given, in which the subject was required to furnish the correct response word for each of the stimulus items. Before the test trial the examiners took pains to remind each subject of the paired-associative task of the previous week. In general, this reminder appeared to be effective.

The *verbal-discrimination task* consisted of 14 pairs of common three-syllable nouns. The subjects were run to a criterion of two successive correct trials or a maximum of eight trials, using the anticipation method. The order of pairs within trials and of words within pairs was determined randomly with two blocks of four trials each, except to ensure that within pairs each word appeared equally often in each position. On a preliminary familiarization, or study, trial, each pair was presented along with information as to which word was correct. On subsequent trials, each pair was presented without that information; the subjects were asked to select the correct items and were then provided feedback by the examiner.

The *word-association test* consisted of 50 Spanish words, most of them equivalents of the English words for which Palermo and Jenkins (1966) obtained word-association data from first- through fourth-grade children in the United States. The general familiarity of the words, even to young unschooled

children, was ensured through extensive pilot testing. The instructions were similar to those used by Palermo and Jenkins, except that somewhat greater stress was placed on speed of responding.

RESULTS

Primary interest was in comparisons of the high-schooling and low-schooling groups' performance on the five tasks described. Those tasks are discussed here in a somewhat different order than that above. It should also be noted that among these comparisons there is considerable variation in the numbers of subjects in each group. (Recall that the groups began with 35 children in each.) Three factors account for this. First, two subjects were lost from each group because subsequent checks showed them to be outside the prescribed age-range. (It turns out that determining the ages of lower-class children in Cali is often extremely difficult.) Two additional subjects were lost from the low-schooling group, and three from the high-schooling group, because their families moved away. Thus, each of the tasks described was administered to 30 high-schooling and 31 low-schooling children.

Verbal-discrimination learning

Figure 1 depicts the performance of the two groups over the eight verbal-discrimination trials. In plotting these curves, subjects who met the criterion of two consecutive correct trials were assigned scores of 14 (perfect scores) for all subsequent trials. That criterion was met by 26 of the 30 subjects in the high-schooling group and by 22 of the 31 subjects in the low-schooling group. The difference between groups in number of trials to criterion did not approach significance [$t(59) = 1.09, p > .20$], despite the somewhat higher mean performance of the high-schooling group on each trial. Further analysis showed that the superiority of the high-schooling group was marginally reliable on trial 1 [$t(59) = 2.07, p = .05$] but that on later trials the differences did not approach significance. The absence of reliable differences on trials 2-8 may be due to a ceiling effect, in that a higher percentage of children in the high-schooling group than in the low-schooling group performed at maximum on each trial, as may be seen in Figure 2. In summary, then, there is some evidence of more rapid verbal-discrimination learning by the high-schooling than by the low-schooling children, but whatever superiority existed does not appear striking.

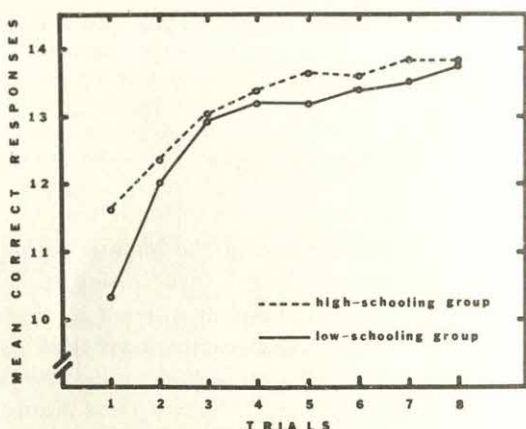


Fig. 1. Mean numbers of correct responses on each verbal-discrimination learning trial

Paired-associate learning

The difference between the groups in paired-associate learning is clear. As shown in the left portion of Figure 3, the high-schooling group showed markedly higher performance on its first, and only, three trials (N-2, N-1, N) than did the low-schooling group on its first three trials (N-5, N-4, N-3), with the difference in mean numbers of correct responses significant on each of these trials [in each case, $t > 2.20$, 59 *df*]. Not until its sixth trial did the low-schooling group reach the high-schooling group's level of performance on its third trial. At that point (on trial N, the last trial for each group), performance for the low-schooling group was very slightly, and nonsignificantly, higher than for the high-schooling group. These differences on the first three trials for each group were found in spite of the same possibility of a ceiling effect as existed for verbal-discrimination learning. In short, evidence for the substantial superiority of the high-schooling children in paired-associate learning appears convincing.

Retention of paired associates

Performance of the groups on the single retention trial given one week after learning is also shown in Figure 3, at the right. The mean drop in performance over the week was 1.80 words for the high-schooling group and 1.68 words for the low-schooling group, a difference that did not approach significance. In fact, looking at Figure 3 as a whole, one

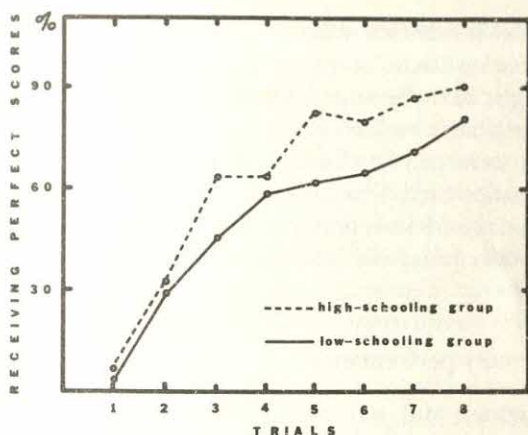


Fig. 2. Percentages of subjects with perfect scores on each verbal-discrimination learning trial

is struck more by the similarity than by the difference in the groups' rates of forgetting; that is, there was no suggestion of any difference whatsoever between the groups in retention once they were equated on original learning.

Note that the paired-associate task a week earlier had been preceded in the same session by the verbal-discrimination task. The question

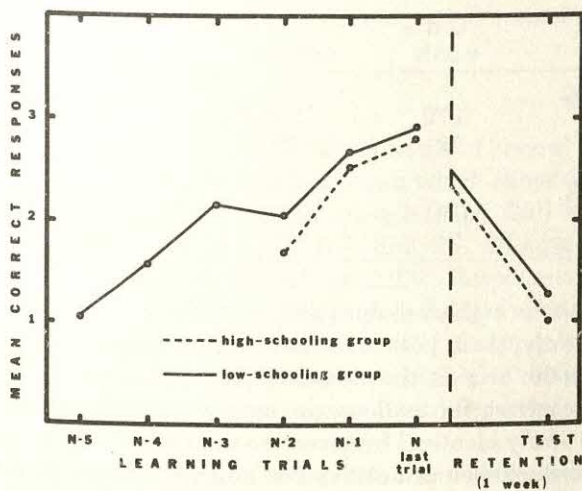


Fig. 3. Mean number of correct responses on each paired-associate learning trial and on the retention trial after one week

raised by this fact is whether substantial proactive interference by the verbal-discrimination items occurred—more importantly, whether that interference might have been greater for the high-schooling group. A check of that possibility yielded negative results on the latter point, for the two groups were nearly identical in terms of intrusions from the verbal-discrimination task. Thus, the data offer no support for the possibility that a basic retention superiority of the high-schooling group was obscured by greater proactive interference from the verbal-discrimination task.

Recognition-memory performance

The mean numbers and percentages of words judged 'old' are shown for each group, and for each category of words in the recognition test, in Table 1. Our major interest was the frequency of 'old' judgments to associates and to acoustically similar words in relation to unrelated words. An inspection of Table 1 shows that such false recognition within the high-schooling group was more than three times as frequent for associates than for unrelated words, and more than twice as frequent for acoustically similar than for unrelated words. Both differences be-

Table 1. Mean numbers and percentages of judgments of 'old' for each category of words by each group

Group	'Old' words	Associates	Acoustically similar words	Unrelated words
High schooling				
Mean	5.79	1.62	1.04	.47
Percent	72	20	13	6
Low schooling				
Mean	5.94	1.65	1.59	1.65
Percent	74	21	20	21

tween means were highly reliable [$t(23) = 4.76$ $p < .01$, and $t(23) = 2.75$, $p < .05$]. Clearly, then, both associative and acoustic factors appeared to play an active role in the encoding process for the high-schooling children. In contrast, for the low-schooling group false recognitions occurred with nearly identical frequency to the three categories of words, suggesting that neither associative nor acoustic factors played an important role for those children.

The frequency of correct recognitions of 'old' words suggests that the

low-schooling group was fully as accurate as the high-schooling group, since the difference between the groups did not approach significance. This is misleading, however, in that false recognition of unrelated words was considerably higher by low-schooling than by high-schooling children [$t(39) = 2.55, p < .05$]. When difference scores are used ('old' words versus unrelated words), the superiority of the high-schooling group was marginally reliable [$t(39) = 1.84, .10 > p > .05$]. Two conclusions are suggested by these results: first, that the high-schooling group was somewhat superior in general recognition accuracy; second, that the low-schooling group employed a more liberal criterion for judgments of 'old.'

It should be noted that the analyses of recognition performance were on 24 subjects for the high-schooling group and 17 for the low-schooling group. These reductions were due to the failure of 20 children to 'catch on' to the procedure. The most common evidence of this failure was seen in a number of children who identified all words in the test list as 'old.' The fact that the subject loss was twice as great for the low-schooling group raises questions of interpretation for the data reported above. Fortunately, it seems likely that this differential loss affected the degree rather than the nature of the difference between groups. That is, one would imagine that the children who were dropped would be among the least advanced in terms of learning and memory abilities, so that the absence of more of these children from the low-schooling than from the high-schooling group would tend to reduce the actual differences in encoding characteristics and in overall recognition accuracy.

Word-association data

The two groups first were compared in terms of several quantitative response characteristics. The frequency with which subjects gave no response was reliably greater for the low-schooling than for the high-schooling group [$t(58) = 2.43, p < .05$]. Reliable differences also were found in response homogeneity and variability. Response variability, or the mean number of different responses per stimulus word, was greater for the low-schooling group [$t = 3.10, p < .02$]. Conversely, occurrence of any of the three most common responses per stimulus word was greater for the high-schooling group, but that difference failed to reach significance.

A second set of findings concerned certain qualitative response characteristics that frequently are examined in analyses of word associations. One of these is the extent to which the responses were superordinates of

the stimulus items (say, the response FRUTA [FRUIT] to the stimulus PIÑA [PINEAPPLE]). Such responses occurred significantly more often among children in the high-schooling group [$t(58) = 2.50, p < .05$]. Next, the groups were compared in frequency of paradigmatic words. In the case of adjectives, the frequency of such responses was reliably greater for the high-schooling group than for the low-schooling group [$t(58) = 2.13, p < .05$]. For nouns and verbs, the differences in this respect were not reliable. A closely related comparison concerned the frequency of responses classified as syntagmatic (a syntagmatic response is one of a form class different from that of the stimulus word and one which reflects syntactic orders; say, the response MONTAÑA [MOUNTAIN] to the stimulus word ALTO [HIGH]). Nor was there a significant difference between the groups in the frequency of responses that were highly similar acoustically to the stimulus items (say, HOMBRE-NOMBRE).

DISCUSSION

In general, the data described are in line with certain expectations derived from current psychological theory and some knowledge of the nature of early school experiences. This is particularly true in the case of associative encoding for memory. The emphasis during the early school years on conceptual and associative relationships among words was expected to result in an increased use of the verbal-associative attribute in encoding. The results were consistent with this prediction, in that false recognitions occurred much more frequently for associates than for unrelated words. The interpretation preferred here is that the stage for such false recognitions was set at the time the first list was presented. That is, when the to-be-remembered words were presented, associates of these words occurred as implicit associative responses. It is assumed that such responses constitute one portion of the information that was encoded for memory during the learning of verbal associations by the high-schooling group. In contrast, there was no such evidence in the case of the low-schooling children: associates were falsely recognized no more frequently than were the control items. Taken together, these data are precisely in line with expectations about the effects of early schooling on such processes.

The results on false recognitions of acoustically similar items were not precisely as predicted. In particular, clear-cut evidence of the importance of the acoustic attribute for the low-schooling children had been anticipated. Such evidence simply was not found. If neither asso-

ciative nor acoustic factors played an important role in the encoding of these children, what factors were important? No answer is provided by these data, and we are left with an intriguing question to tackle in the future.

For each of the three learning tasks employed, there was evidence of better performance by the high-schooling children. In the case of paired-associate learning, this superiority was impressive. This difference may have been due, at least in part, to a failure of the low-schooling subjects to use the verbal-associative and the acoustic attributes to encode the material for memory. In the case of verbal-discrimination learning, only a small and marginally reliable difference was found. The lack of a substantial difference here does not seem surprising in view of the apparently different process underlying verbal-discrimination learning, namely, a heavy dependence on the frequency attribute. The superiority of recognition memory in the high-schooling group took the form of a lower frequency of false positives for that group, suggesting that their criterion for a judgment of 'old' may have been more stringent.

The retention data suggest that the inferiority of the low-schooling children with the paired-associate task was confined entirely to the learning (encoding) stage. Once their level of learning was brought up to that of the high-schooling group, the low-schooling children retained at least as much. But if the high-schooling children did have a richer learning history, why wasn't there greater proactive interference, and thus greater forgetting, in their case? Again we can only speculate. One possibility concerns the nature of the material used—highly familiar nouns paired with common boys' names. Perhaps the two groups simply did not differ in their stores of the *particular type* of information that would provide interference in this case.

The word-association data provides further evidence that the two groups differed substantially in verbal development. In general, the difference found between the two groups was similar to those that have been reported for older (nine- or ten-year-old) compared to younger (six- or seven-year-old) children (Cramer, 1969, pp. 141-151). Note that in such previous age comparisons the groups differed both in amount of schooling and in other age-related differences. In the present case we see roughly the same sorts of differences in word associations with age held constant.

In what ways did the two groups differ other than in their formal educational experiences? Unfortunately, that question cannot be answered in any complete sense, so that causal conclusions about the effects of schooling would be on shaky grounds. However, several factors suggest

that the differences found were due in large measure to the schooling variable. One of these is that all of the subjects were drawn from a relatively homogeneous lower-class community. That is, there were no obvious socioeconomic differences between the two groups. Another factor is that among those low-schooling children who had been in school for more than one year, the frequency with which they had been retained in first grade was lower than for the high-schooling group. A third factor is that it is quite common in Colombia for children to be held out of school, because there is no school nearby or because of family demands on even quite young children. In other words, there are plausible alternatives to the hypothesis that the low-schooling children began school later because of intellectual deficits.

The present study was seen as an initial step in discovering causal relationship between specific early educational experiences, on the one hand, and specific process changes, on the other. The results were encouraging enough to suggest going on to research (incorporating the necessary experimental controls) on the effects of various instructional inputs to three- through six-year-old children. This, in turn, must be preceded by the development of effective techniques to assess learning and memory processes in children of this age. Preliminary work in this direction is underway.

Notes

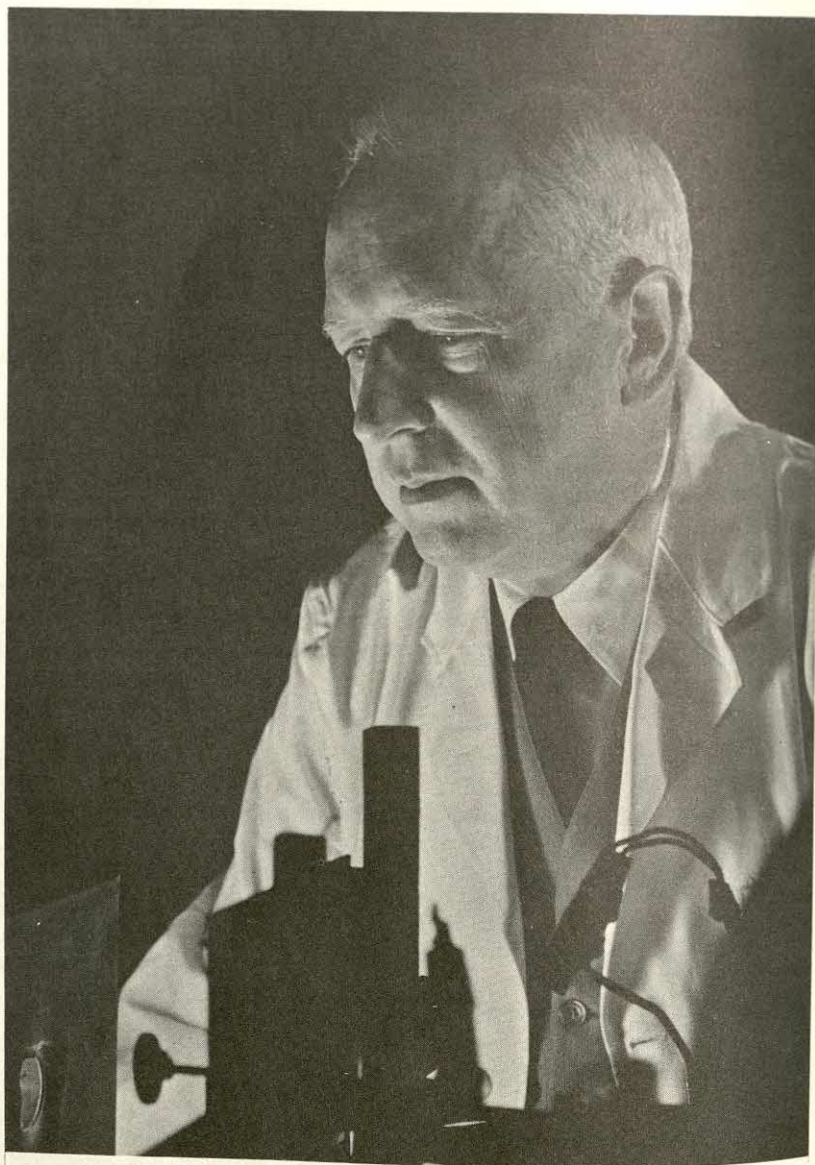
This project was supported by a grant from the Council for Intersocietal Studies, Northwestern University. I am grateful to the Director of the Council, Richard Schwartz, and to B. J. Chandler, Dean of the School of Education, all at Northwestern, for their support. I am especially indebted to Angela Rojas and Ruth Montague, whose assistance at various stages of the project was invaluable. Portions of this paper were presented at the 1970 meetings of AERA, Minneapolis, Minnesota. Received for publication May 7, 1971.

1. Drs. Leonardo Sinisterra, Harrison McKay, and Arlene McKay, and their staff at the Ecological-Nutritional Field Station in Cali provided assistance and cooperation without which this project could not have been completed.

References

- Bach, M. J., and Underwood, B. J. 1970. Developmental changes in memory attributes. *Journal of Educational Psychology* 61:292-296.
Cramer, P. 1968. *Word association*. New York: Academic Press.
Ekstrand, B. R., Wallace, W. P., and Underwood, B. J. 1966. A frequency

- theory of verbal discrimination learning. *Psychological Review* 73: 566-578.
- Hall, J. W., and Crown, I. 1970. Associative encoding of words in sentences. *Journal of Verbal Learning and Verbal Behavior* 9:303-307.
- Hall, J. W., and Ware, W. B. 1968. Implicit associative responses and false recognition by young children. *Journal of Experimental Child Psychology* 6:52-60.
- Palermo, D. S., and Jenkins, J. J. 1966. Oral word association norms for children in grades one through four. Pennsylvania State University Research Bulletin 60.
- Underwood, B. J. 1969. Attributes of memory. *Psychological Review* 76: 559-573.
- Underwood, B. J., and Freund, J. S. 1968. Errors in recognition learning and retention. *Journal of Experimental Psychology* 78:55-63.



Clarence H. Graham

NOTES AND DISCUSSION

CLARENCE HENRY GRAHAM: 1906-1971

With the death of Clarence Henry Graham, professor of psychology at Columbia University, on July 25, 1971, American and indeed international psychology lost an outstanding figure and graduate education a skilled and greatly loved teacher.

Born on January 6, 1906, in Worcester, Massachusetts, Graham was a first-generation American. His mother and father, who met and were married in this country in 1903, were Protestants from county Donegal, Ireland. A great-aunt Maggie, herself childless, was the dominant figure in the household, and plans for young Clarence's education and social life seem to have emanated from her rather than from his more permissive parents. There was a younger sister, who predeceased him, and two younger brothers.

Viewed in the light of current population mobility, Graham's education appears somewhat limited, but it was unquestionably efficient. He entered the public school system in Worcester at the age of five (kindergarten), continued at Cambridge Street School (1911-1919), went on to South High School (1919-1923), and entered Clark University in the fall of 1923. All three degrees were from Clark (A.B., 1927; A.M., 1928; Ph.D., 1930). Thus, for a continuous span of 21 years he received his formal education in Worcester institutions.

As a first-year student at Clark, Graham was more than a little popular and was elected president of his class. Those who remember him as an undergraduate recall his incessant banter and clowning on the baseball field and his good-natured management of the soccer team. He was possessed of an irrepressible good humor. A fraternity brother writes of him: "His wit was original, dry, droll, and spontaneous. . . . Everyone who knew him loved him, and, as he grew in stature and maturity, respected him." Graham began by majoring in chemistry, changed to literature, and ended up in psychology after coming under the influence of one of Clark's great teachers, John Paul Nafe. In his final undergraduate year he was editor-in-chief of the college literary magazine and was elected to the senior honorary society, Gryphon.

Clarence Graham's professional career centered around three institutions: Clark, Brown, and Columbia. His preparation and some of his early teaching

were at Clark, where in graduate studies he was influenced most by Nafe and Hunter. He admired the former, a 'regular fellow' who always made his students seem important, more as a firm friend than a mentor, and the latter for his consistently objective and analytical outlook on psychological problems. Graham was to remain close to both as long as they all lived.

Largely as the result of rapid staff changes at Clark in the early 1930s, he returned from Philadelphia after two highly productive research years with Hartline and Granit at Pennsylvania (during the first of which he also taught at Temple) to join the Clark faculty as an assistant professor. He took over responsibility for experimental psychology, Nafe having departed for the chairmanship at Washington University, Saint Louis.

The Brown years were 1936 to 1945. Hunter had replaced Carmichael there when the latter moved to Rochester as dean of the faculty, and Hunter needed Graham. He was rapidly advanced to associate professor in 1937, and to professor in 1941. The call to Columbia, occasioned by the retirement of Robert S. Woodworth, came at the end of World War II and brought with it a flood of graduate students in the postwar buildup. Graham found himself responsible, at any one time over the next half-dozen years, for eight to ten graduate theses. The total number of doctoral candidates he 'saw through' in his nine years at Brown and his quarter of a century (plus one) at Columbia was close to 70.

Graham's research interests were, in one sense, varied; in another, they were remarkably restricted and intensive. Variation seems the theme if one considers the range of topics embraced: from visual contrast to 'emotionality' in rats; from atmospheric 'boil' to nerve impulses in the optic nerve of *Limulus*; from the design and calibration of an animal runway to monocular movement parallax. At the same time, with only a few notable exceptions, his work was systematic and was concentrated on the visual process, mainly on color vision in his more mature years, but with increasing involvement in problems of space and movement perception.

The Clark and early Brown work had largely to do with temporal, intensive, and spatial relations in achromatic vision, much of it revolving around the basic problem of intensity discrimination. The interests of the later Brown years were those dictated by wartime exigencies, especially those of the Applied Psychology Panel of the National Defense Research Committee, on which Graham served with conspicuous devotion. He turned to visual problems connected with gunfire control, more particularly those of binocular range finding and height finding.

Curiously enough, he also got involved with the very specific problem of personnel selection and classification arising from the Navy's desire to exclude from training for submariners those whose emotional stability could be questioned. The personal inventory he and Walter Shipley devised for this purpose seems an oddity in context with his other scientific and technological interests. Perhaps, as he himself said of it, the inventory work may have owed its origin to his psychoanalysis in the late 1930s. However that may be, there are those who think that a visual psychophysiologist who works on an intimate human problem is not necessarily out of character—if his name is Clarence Graham. Those who knew him know also that he was a most sympathetic and understanding person, basically a humanitarian.

Lorin Riggs, Graham's most distinguished student and for many years his colleague and research associate, tells of an incident during their work in the mid-1930s on the spectral-visibility curve of the white rat. To record the animal's electroretinogram, or ERG, it was necessary to strap the rat down and keep its eye rigidly aligned with the stimulus and the wick electrode that contacted the cornea. Throughout the experiment Graham would reopen the chamber and tug on the restraining tape to ensure immobility, but never without repeating over and over, "Poor damned animal, poor *damned* animal." As Riggs says, "His compassion was easy to understand if you knew Clarence."

The move to Columbia in 1945 brought with it not so much a new set of research interests (for Graham was to remain a considerable systematist in visual science) as a distinct swerving of his trajectory toward a preoccupation with problems of color vision. To an incalculable degree this must have been influenced by Selig Hecht, whom he had long admired, and by one of Hecht's postdoctoral fellows, Yun Hsia. Graham and Hecht gave a seminar together in 1946-1947, an association precipitously terminated by Hecht's shockingly untimely death at the end of that year. Graham inherited much of the equipment of Hecht's laboratory and Hsia came over to work with him; the two constituted a team that was to perform effectively together for 21 years, until Hsia's death in 1968.

Over this period, color was the central theme. Graham and Hsia first did a series of studies on color-normals, finding the spectral-brightness maximum to be approximately where the ICI 'standard observer' put it, but finding additionally that by relieving the measurement problem of relative-luminosity coefficients and ascertaining absolute sensitivities there showed up marked depressions in the overall spectral-luminosity curve in the neighborhood of 460 nm (blue) and 600 nm (red). The ICI curve also underestimates luminosity in the short waves, they discovered. Then, turning to dichromia, they found that whereas protanopes had the familiar losses at the long-wave end of the spectrum, deuteranopia was also characterized by a reduction of luminosity in the green, a point on which there had been little really trustworthy data.

If a particular problem were to be singled out as representing Graham's research focus at the height of his career, it would have to be the intensive study, with Hsia, of a remarkable case of unocular color blindness discovered, almost by accident, in a psychology class at Barnard. Over the years 1958-1962, nearly a dozen papers were devoted to the details of this young woman's vision. Never before had there been such a thorough work-up of such a case. There came under careful measurement seven crucial functions: threshold spectral luminosity, critical fusion frequency, binocular brightness equations, complete color-mixture curves, wavelength-discrimination data, color naming, and binocular chromatic matches. The rare opportunity of having a color-normal eye check on the characteristics of the color-blind one was taken full advantage of, and there resulted not only a set of highly precise measurements but real advances in color theory as well.

Two other extraordinary opportunities to study deviant color phenomena presented themselves in the middle and late 1960s. One concerned a colleague, Professor Stephan, who, by reason of a vascular accident, had a parachromasia covering the central 18 deg of one eye. By measurement of lu-

minosity functions and neutral-band locus, supplemented by binocular hue matches and color naming, it was established that the defect was not far from that given in the classical picture of congenital tritanopia. The study was especially valuable for its information on color naming in such a visual system under widely varying illuminations. The other case was that of a young woman previously judged by competent authority to be totally color-blind. Graham and Hsia were able to demonstrate some remaining, if rudimentary, foveal vision that yielded a little color sensitivity. Thus there was added to the already complicated catalog of color deficiencies yet another variant on the main theme.

Graham's career was not spent exclusively in the laboratory. He became well known in the psychological and physiological circles of Europe by reason of his effective year with the London Branch of the Office of Naval Research (1952-1953). He also made two extended visits to the Far East, giving a seven-week seminar at Kyoto in the spring of 1952 and, eight years later, traveling to twelve of the most active and prestigious Japanese institutions. At each he lectured, talked with laboratory workers, and assembled extensive notes that were subsequently the basis of two publications on psychology in Japan. He also returned to England for a year's stay, working in London at Professor Wright's Imperial College laboratory. This program was carried out under a Guggenheim fellowship.

Honors came to Clarence Graham in some abundance, beginning with the Society of Experimental Psychologists' award of the Warren Medal in 1941. Other such distinctions were the Tillyer Medal of the Optical Society of America (1963) and a Distinguished Scientific Contribution Award from the American Psychological Association in 1966. He was elected to the National Academy of Sciences, the American Philosophical Society, and the American Academy of Arts and Sciences. Brown University awarded him the honorary degree of Doctor of Science in 1958, on the occasion of the dedication of the Walter S. Hunter Psychological Laboratory. His war work was recognized by the Presidential Certificate of Merit in 1948.

Publications there were in goodly number. At the time he was awarded the Tillyer Medal, there was assembled a bibliography of 77 titles. Since then the output has gone up to 90, including one that eclipses all others: *Vision and Visual Perception* (Wiley, 1965). This magnificent volume, completed with the aid of five well-known colleagues (Brown, Bartlett, Hsia, Mueller, and Riggs) was many years in the making. The original agreement, in the postwar years, called for a single-author book, but it became apparent that any really complete and systematic treatment of the burgeoning topic of vision would have to be a joint enterprise. Graham himself contributed 9 of the 20 chapters, collaborated on another, and edited the entire volume. It will long remain a tremendously successful monument to him.

Frank A. Geldard, *Princeton University*

GÖSTA EKMAN: 1920-1971

While deeply engaged in productive research, Gösta Ekman was stricken with illness in 1969, and never recovered. He died on May 5, 1971, at 50 years of age.

Gösta Ekman was born on June 3, 1920, in Lund, where he pursued the greater part of his university studies. After studying medicine for a brief period and taking a baccalaureate in literature, he turned his attention to psychology. In 1947 he became doctor of philosophy with a thesis on reliability and constancy. He succeeded David Katz as professor of psychology and chairman of the Psychological Laboratories at the University of Stockholm in 1952. He held these posts until 1967, when he was awarded a personal research professorship in psychophysics and scaling methodology by the Swedish Council for Social Science Research. His research work was performed at the laboratories, where he also continued his work with graduate students.

For 15 years Gösta Ekman was head of the University of Stockholm Psychological Laboratories, which during this period became one of the leading research institutes in Sweden and gained an international reputation. Though this entailed a great amount of work, he was able to perform research of an unusually high quality. His early work was concerned mainly with differential psychology and with testing-theory problems. In addition to his doctoral thesis on reliability and constancy, his main work during this period includes a book on differential psychology (1952), a lucid presentation of problems of method in differential psychology. From the mid-1950s his research was concerned almost exclusively with modern psychophysics. He made new contributions to psychophysical scaling methods, but his most important contributions were perhaps his content model in multidimensional scaling and the experimental work leading up to what S. S. Stevens has called Ekman's law, "the subjective counterpart of Weber's law." This law states that variability measured in psychological units is linearly related to psychological magnitude measured in the same units.

In spite of his great administrative responsibilities as chairman of the rapidly expanding laboratories and as adviser to a great number of students, Gösta Ekman always found time for research. This was characterized by originality, versatility, and methodical skill, and typical of his presentation were analytical preciseness, lucidity, and elegance. These characteristics, combined with great personal generosity, attracted many students, not only Swedes, to his laboratory; many of them now occupy important positions in psychology.

By his activities as research worker, teacher, and head of the laboratories and by his work in research councils and such, Gösta Ekman came to mean more for the development of psychology in Sweden than anyone else. Internationally he stands among those who have made lasting contributions to modern psychophysics and psychological measurement.

David Magnusson, *University of Stockholm*

ANNOUNCEMENT

NSF Chautauqua-type Short Courses for College Teachers

The American Association for the Advancement of Science has announced that the third annual program of NSF Chautauqua-type Short Courses will begin in October 1972. There will be places for 3,000 participants, with ten courses being offered at each of 12 Field Centers located throughout the country.

Topics cover a wide range of fields in the natural and social sciences, mathematics, and engineering. Most of the courses are interdisciplinary in nature, including the applications of mathematics to college teaching.

The courses consist of two days of intensive classwork at one of the centers, followed by approximately three months of independent research or study, and ending with another two-day session for discussion of the work performed and general synopsis.

The announcement brochure of the 1972-1973 program can be obtained by writing to NSF C-T Courses, AAAS, 1515 Massachusetts Avenue, N.W., Washington, D.C. 20005.

CORRECTION

The review of *The Perception of People and Events* in the December 1971 issue of this journal (Vol. 84, No. 4, pp. 591-593) was by Phillip Shaver, of Columbia University.

BOOK REVIEWS

Julian Hochberg
Columbia University

Cognition and Affect

Edited by John H. Antrobus. Boston: Little, Brown and Co., 1970. Pp. x, 210. \$6.95.

Some of the oldest problems in psychology are addressed in this book: 'mental processes revealed by reaction-time experiments,' recall of affect-laden material from early childhood, daydreaming, visual imagery, and the relationship between attitudes and behavior. The rampant enthusiasm for contemporary cognitive psychology, here said to be undergoing a renaissance (see the introduction by Antrobus) and in several recent articles hailed as revolutionary, actually seems justified by some indications in this book. Of course, the relationships between affective and cognitive processes were never quite as neglected as we sometimes hear these days, when behaviorism is assaulted as if it were an oppressive ruler. But the powerful methods developed during the reign of strict behaviorism have now been combined with theoretical constructs borrowed from the sciences of communication and artificial intelligence to yield a robust cognitive psychology. Perhaps we can now begin extending it to include the notoriously uncomputerlike emotions.

Cognition and Affect, which resulted from a conference at the Center for Research on Cognition and Affect of the City University of New York, was designed to represent the diversity of contemporary cognitive psychology. By accomplishing this in roughly 200 pages, it necessarily fails to maintain any semblance of coherence. Still, the individual chapters are of high quality.

Saul Sternberg reviews his excellent studies of retrieval and high-speed scanning of information in memory. Some professionals will already be familiar with his work, which has been described in several journal articles; but nonexperts will appreciate the clear account of Donders's (1868) 'subtraction method,' based on the notion that the temporal interval between stimulus and response can be regarded as the summed durations of several information-processing stages. By altering some of Donders's assumptions and devising ingenious new procedures, Sternberg has been able to draw a number of striking conclusions about memory. And although he says nothing at all about

affect, the methods Sternberg has developed will also prove useful to researchers concerned with the relationships between affect and cognition.

An example of such methodological utility is provided in Jerome Singer's chapter on daydreaming. Using techniques developed for studies of vigilance and attention, he and his colleagues have been able to reveal the remarkable pervasiveness of fantasy and to infer some of its lesser-known functions. Regarding pervasiveness, Singer concludes—perhaps incorrectly—that daydreaming can occur in parallel with processing of information from the external world. It seems more likely, considering both his evidence and recent research on central processing capacity, that daydreaming is compatible only with reception of highly predictable or redundant signals. Still, much of our everyday world is highly familiar, predictable, and redundant; Singer's questionnaire and experimental studies leave little doubt that we often take advantage of this fact to occupy ourselves with the exploration of fantasy worlds. Regarding function, Singer claims that in addition to the occasional cathartic effects of daydreaming, the "adaptive function of the ongoing stream of thought, and of nocturnal dreams as well, involves . . . organization and attempted integration of the major unfinished business in our lives" (p. 152).

Singer's results are themselves highly compatible with Ulric Neisser's well-known 'constructionist' approach to cognitive processes. In the present volume Neisser elaborates the account of visual imagery given in his 1967 book, proposing that similar constructive processes generate both percepts and images in the visual system. Several recent studies have demonstrated, in support of this view, that attention to visual but not auditory stimuli interferes with visual imagery. Neisser also reviews research on individual differences in the experience and use of imagery, and he concludes that differences in reported vividness do not seem to have functional significance, whereas differences in spatial-reasoning (visualization) ability certainly do. Unfortunately, Neisser does not discuss how these important individual differences develop or whether they are related to other perceptual abilities, even though his theoretical approach would suggest common mechanisms and origins.

Silvan Tomkins, whose earlier work contributed to the establishment of the Research Center for Cognition and Affect, here offers a condensed but still long and rather rambling account of his theory. His ideas seem virtually to erupt from his mind, with little regard for empirical evidence and few references to the work of other psychologists. (Whereas Sternberg cites 50 papers in about 40 pages, Tomkins cites only 7 in his 70 pages, which make up nearly one-third of the book.) There are many undocumented generalizations; for example, that "the information to be remembered must be transformed into a compressed, miniaturized equivalent form" (p. 70), which, Tomkins says, is accomplished by faster and faster covert repetition. "In addition to speed, the sounds must be clipped and abbreviated without destroying their essential message" (p. 73). Yet, even if the lack of evidence is frustrating, it must be admitted that Tomkins is brilliantly provocative. In ranging over topics as diverse as memorization of telephone numbers, reactions to novel and familiar stimuli, and therapeutic evocation of childhood memories and behavior patterns, Tomkins offers creative proposals that are eminently researchable.

In the final chapter, Milton Rosenberg attempts to distinguish between two brands of cognitive-consistency theory, Festinger's and his own. The latter, Rosenberg claims, explains more 'real world' social behavior and is also more

humane, because it portrays human action as the result of beliefs, attitudes, and decisions, whereas Festinger's theory would make behavior ("inauthentic action," according to Rosenberg) the determinant of attitudes and beliefs. As for the truth of the respective theories and their range of application, no firm conclusion is possible at present. Moreover, any complete evaluation of research in this area must now include a discussion of attribution theory and perhaps also of the impression-management theory outlined by Tedeschi, Schlenker, and Bonoma in a recent *American Psychologist*. As for the relative humaneness of the various theories, it is interesting to note that many social psychologists thought dissonance theory humane because it referred to 'cognitions,' and these seemed more characteristic of humans than did 'stimuli and responses.' Tedeschi et al. explicitly claim that their theory is more humane than Festinger's because it characterizes people as active strategizers rather than passive victims of experimentally manipulated dissonant cognitions. Yet their notion that subjects lie and deceive in order to save face and gain approval from an experimenter is far from most versions of the humanistic ideal. We must finally admit that hypotheses cannot be validated by arguments about humanism.

Much of the material in this book is available elsewhere, and the separate chapters do not gain a great deal from their joint presentation. Although the book resulted from a conference, it contains no record of the exciting discussions that must have taken place. It could nevertheless serve as a good introduction for both graduate and undergraduate students to some of the most challenging problems in contemporary psychology. This is surely a worthy purpose, and it is fortunate that the book has been made available in paperback as well as hardcover.

Phillip Shaver, *Columbia University*

Behavior and Conscious Experience: A Conceptual Analysis

By Kendon Smith. Athens, Ohio: Ohio University Press, 1969. Pp. 140. \$6.00.

This book is not addressed, as its title suggests, simply to concepts in use and their meanings. The author is after what we really mean by the concepts of stimulus, response, learning, and conscious experience—after no less than "the fundamental nature of the object or phenomenon to which [each of these terms] is customarily applied" (p. 3). The analyses proceed largely by way of found meanings, in biology, psychology, and everyday thought. But, as his aim dictates, the author tries to push beyond the literature in order to get better hold of the phenomena underlying the concepts that are of interest. He evidently finds his definitions of those concepts an improvement over what we had, or did not quite have, before.

Learning theorists often refer in their discussions of the response concept to activities of the whole organism, molar behaviors, actions, doings, and the like. The author takes such references figuratively, as symbolic of a repressed criterion. What is meant though unsaid, he believes, is that a response is that kind of liberation of chemical energy by a living object that is *produced by conscious experience*. No doubt Watson and others would resist this claim. Presumably Smith would counter with "the evident impossibility of understanding [responses] on any basis other than the one [he has himself] sug-

gested" (p. 42). In the same contexts some of the theorists refer to muscle twitches, movements, and simple responses. These phenomena Smith brushes aside, saying that no psychologist as such would be interested in them unless they were sensitive to conscious experience. He finds "strong support" in N. E. Miller's definition of response: "any activity by or within the individual which can become functionally connected with an antecedent event through learning." Interpreting 'antecedent event' as necessarily a cue, Smith reasons that since a cue traditionally is something that evokes conscious experience, Miller is "almost explicitly concurring" with his own definition of response. That the behaviorist use of 'response' implies the notion of the subject's awareness is an intriguing suggestion that would have some influence were it demonstrated. But it takes more to accomplish this than reading in symbolic criteria. It takes a detailed analysis of a variety of actual uses of the concept, in both observational and explanatory contexts, to determine whether an experiential dimension is essential to the concept's function in the behavioristic framework.

For a number of learning theorists, a stimulus is anything that produces a response. From this, using his definition of response, Smith readily jumps to a definition of a stimulus as *anything that produces a conscious experience* (which concept he modifies later but which remains basically the same). He finds support in Watson's mention of energies that effect visual reactions or auditory stimulation, events customarily associated with sensation at the time that Watson was writing. Smith quickly concludes, "We realize at once that even Watson had not forsworn the notion of stimuli as instigators of conscious experience" (p. 48). Perhaps not, but is this a way to show it? And given the shaky grounds on which Smith's own definition of response was built, should he not do more to justify this second 'extraction' from the behaviorists?

The contemporary concept of learning, says Smith, basically involves connection of a new response to a stimulus, the process depending on events which are themselves stimuli. In Smith's hands, the behaviorist definition of learning turns into "the attachment of a new response to a conscious experience, by conscious experience" (p. 76). The author finds the justification he clearly needs for this new attribution in a certain resistance that psychologists manifest, namely, an unwillingness to grant to systems presumed to lack consciousness the ability to learn. One-celled animals are among such systems. He dismisses as source for this resistance the view that one-celled organisms might be too simple to learn, since "it is quite possible that the structure of a single cell provides all the elements necessary for such a process" (p. 77). That is possible—and it is also understandable that psychologists who thought of learning in terms of connections between cells would doubt evidence of learning in one-celled animals. When evidence for spinal conditioning first appeared, Smith says, there were no immediate and widespread attempts to replicate those results, nor was there even any great interest in them. He explains this by saying that psychologists were unable to conceive of a conscious spinal cord. He does not mention the possibility that spinal conditioning is consistent with the widely held concept of learning as the attachment of responses to stimuli, and for that reason would not be greeted with excitement. Whatever the final value of Smith's definition of learning, the case he presents for it, either as an abstraction from the literature or as an insightful contribution, is not in itself persuasive.

Conscious experience, according to Smith, refers to sensations, perceptions, images, thoughts, dreams, wishes, and emotions. It is a process, a flow of events that are (emphatically) physical, "changes in the very substance of the brain" (p. 12). Starting from the implicit everyday notion of consciousness, he arrives, by a series of approximations, at a definition of conscious experience as "an internal event to which an arbitrary response can be attached directly, by a process of following that response, when it occurs to that event, with a prespecified set of circumstances" (p. 119). On what grounds is this definition to be taken seriously, as capturing the distinguishing characteristic of conscious experience? The grounds are direct intuition when there is at first some doubt as to whether the system under study is capable of conscious experience and yet the conditions specified in the definition are met.

Smith illustrates his definition with a hypothetical instance: the repeated instrumental conditioning, with varying response, of a spinal preparation. If this were truly accomplished, "the investigator would have to concede sentience to the spinal cord . . . [because] only by assuming that the shock on the tail was in truth *felt* could the experimenter account for the existence of such a strikingly adaptive response to that shock" (p. 116). Now, there is no question that we frequently use 'conscious experience' and other mentalistic concepts to account for behavior and changes in behavior. Taking note of this explanatory function, however, merely scratches the surface of what conscious experience is. The definition says nothing about the event itself except to mention one way in which it might relate to behavior. And certainly it does nothing to render the rich concept on which we commonly rely. For example, at the heart of the everyday notion of conscious experience is its having a content, however disparate from physical reality; that is to say, conscious experience in various of its forms is about something beyond itself, is intentional.

A sense of incompleteness may have motivated what the author does late in his last chapter. He tries to make the picture he has drawn more meaningful and understandable by espousing a "motor theory of consciousness." The theory is not quite that, for conscious experience is said to occur with activation of the sensory projection areas by whatever means. Perception is a form of behavior, a "tentative, covert, 'identifying response'" which elicits conscious experience by feedback to projection areas. Similarly for images and thoughts, wishes and emotions: "All 'conscious phenomena' reduce to mosaics of sensation. . . ." In view of his original goal of getting at the fundamental nature of the objects or phenomena referred to by the concepts of interest, and in view of this last hypothesis on the nature of consciousness, Smith might better have addressed his book to a detailed exposition and evaluation of this sensationist theory of consciousness.

Thomas Natsoulas, *University of California, Davis*

Mechanisms of Animal Discrimination Learning

By N. S. Sutherland and N. J. Mackintosh. New York and London: Academic Press, 1971. Pp. xi, 559. \$18.50.

I have mixed feelings about this book, which provides a useful introduction to a large and difficult literature but is marred by what seems to me to be an

undue preoccupation in its pages with the merits of a rather unimpressive theory.

The starting point of the theory is the old (noncontinuity) assumption of Lashley and Krechevsky, long criticized by Spence, that discriminative learning involves selection of stimuli as well as selection of responses. Responses are said to become 'attached,' not to sensory inputs, but to the 'outputs' of dimensional 'analyzers,' whose function is hypothetical (in Krechevsky's sense) rather than perceptual. "To say that an animal has not switched-in a brightness analyzer is not to say that brightness differences are not detected," Sutherland and Mackintosh write, but "only that they do not control choice behavior, and that information about their correlation with reinforcing events is not stored" (pp. 126-127). Performance on any trial is determined by the response attachments of the strongest analyzer and any others of relatively high strength. (The assumption is made that more than one analyzer may be 'switched-in' at the same time.) Analyzer strength also determines the degree to which response attachments are strengthened by reward and weakened by punishment. (There is no provision for classical conditioning.)

An analyzer is strengthened "when its outputs consistently make correct predictions about . . . trial outcomes" (p. 39). The change in analyzer strength on any trial varies directly with amount of reward and also with the "strength of the expectancy confirmed" or "disconfirmed" (pp. 44, 257), although reinforcing the "weaker" of two response attachments in an experiment on probability learning "is assumed to weaken the analyzer" (p. 428). When one analyzer is strengthened, the total strength of other analyzers is weakened correspondingly, but other analyzers in the same modality may be strengthened. When no analyzer makes consistently correct predictions, all analyzers revert to basal levels of strength that depend on genetic factors and on experimental conditions, being greater, for example, when stimuli vary within trials rather than between trials, and increasing with the "discriminability" of the stimuli (p. 132). The authors also assume that analyzer strength may be "conditioned" to "outputs from other analyzers" (p. 163), or even to the "memory of reward or nonreward" (p. 268), and that it is susceptible to interference-produced forgetting.

This, then, is the loose assortment of admittedly vague principles, proposed with varying degrees of conviction, from which Sutherland and Mackintosh attempt to derive the results of a wide range of experiments on blocking and overshadowing, acquired distinctiveness, transfer along a continuum, generalization, reversal learning, dimensional shifting, partial reinforcement, and probability learning. The fit is not very satisfactory, and the critical reader, no matter how sympathetic, soon must begin to feel that a fresh approach is required, regretting all the more that the focus of the book is on the theory rather than on the experimental literature. The literature itself leaves much to be desired, and I have the impression that the theoretical bias of the authors has led them to overlook certain of its obvious shortcomings. At the outset of the book, for example, three experiments are described that are said to demonstrate "reference phenomena" (p. 45) "difficult, although not impossible, for S-R theory to explain" but "readily explained in terms of selective attention" (p. 30). It is significant that the design of each has a serious flaw.

In a chapter on comparative psychology, the authors write convincingly of

the importance of the work (however unconvincing their interpretation of the results thus far obtained), but there is a methodological problem they do not see clearly. To control for the effects of sensory, motor, and motivational variables which cannot be equated across species, I have developed the method of systematic variation. "An alternative approach," write Sutherland and Mackintosh, "is to utilize theoretical understanding of the situations investigated" (p. 432). Their idea is that if a certain difference in the discriminative performance of two animal groups can be attributed to a difference, say, in attention, and if the attentional interpretation leads to the prediction (experimentally verified) of a second difference in the discriminative performance of the same animal groups, then the possibility that both differences in performance are due, not to a difference in attention, but, say, to a difference in drive is ruled out. It should be evident that only variation in drive can tell us whether a difference in the performance of two species is invariant with drive.

M. E. Bitterman, *Laboratory of Sensory Sciences, University of Hawaii*

Stimulus Sampling Theory

Edited by E. D. Neimark and W. K. Estes. San Francisco: Holden-Day, 1967. Pp. xvi, 681.

This book of readings should serve as an excellent textbook for advanced undergraduates and graduate students in a course on the development of mathematical learning theory. It is not all-inclusive in its coverage of that theory, but the authors make no attempt or claim to cover the entire field. The approach is largely that of the students and colleagues of Estes and Bush and Mosteller. The emphasis is on the psychological processes in human and animal learning rather than on the mathematical aspects of theorizing. The book is divided into five chapters covering the development of the general aspects of stimulus-sampling theory, associative learning, probability learning, and discrimination and generalization viewed as single and multiple processes. At the beginning of each chapter is a concise and very informative section that provides background material and a summary of the developments found in the readings.

I think the book is important for a number of reasons, but here I wish to dwell only on one. Stimulus-sampling theory is, to a large extent, a quantitative formulation of the S-R associative tradition that has dominated theories of learning. Most of the readings reflect this underlying theme of associative connections between environmental events (stimuli) and behavior (responses). It is of more than passing interest that only four years after the publication of this book (which covers approximately the prior 15-year period), many of the authors have now taken a cognitive stance and have interwoven concepts of memory into learning theory. Many of the areas covered in the book are not being investigated as much now. The reason for this is not that these areas are now well understood. Rather, as one progresses through the readings one readily appreciates the theoretical problems encountered there because in many instances they still exist today.

I think this book may come to serve as a historical marker of a transition in concepts of learning theory. The zeitgeist today is more cognitive—an em-

phasis on notions of storage, retrieval, and information processing rather than on S-R associative connections. In some of the readings, suggestions of these new notions can be detected. Whether we are simply replacing old terms with new ones or whether these terms shall come to represent important concepts and reflect underlying processes are questions that must be left to future historians.

Donald Robbins, *Emory University*

Le temps psychologique

By J. C. Levy. Science Poche, 16. Paris: Dunod, 1969. Pp. 106.

It took me a long time to read this very little book (less than 25,000 words). As a matter of fact, I read it three times over and still do not feel enlightened about psychological time. It therefore apparently does not serve the intended goal of the Science Poche series, which is, according to the text on the back cover, to offer concise presentations in a "careful didactical way" to assist the overloaded scientist who nowadays needs a permanent education.

The volume deals with the simultaneity and succession of events and with the way the human organism might deal with such events. It promises to offer a rigorous (mathematical) language to discuss the psychological problems inherent in the area of time. This language is the framework of algebraic semi-groups and lattices that makes it possible to concatenate and (partly) order series of events. It seems quite plausible to use this formalism for the description of temporal phenomena, including memory and forgetting, but this is hardly clear from the text, which offers only some vague and anecdotal arguments.

The second part is devoted to a description of (hypothetical) networks that can recognize patterns and which are said to parallel certain psychological processes. Apart from a very loose connection with psychological time, this part is also suffering from unnecessary complexities and confusing arguments. Altogether, not a very illuminating or helpful text.

J. A. Michon, *Institute for Perception, RVO-TNO,
Soesterberg, the Netherlands*

The Dyslexic Child

By Macdonald Critchley. 3d ed. Springfield, Ill.: Charles C Thomas, 1970. Pp. xv, 138. \$6.50.

The Dyslexic Child deals with the reading disorder known as specific dyslexia. The author is very careful to differentiate specific dyslexia from other reading disorders. He bases his conclusions about specific dyslexia on his own clinical observations as well as on extensive reference to the literature. Although Critchley does give precedence to his own observations, he also provides a thoughtful criticism of the reading-disorder literature he surveys.

Critchley defines specific dyslexia as a reading failure with a purely endogenous origin and virtually no associated symptoms. He argues that factors such as the language to be learned and the manner of instruction are not

causally linked to specific dyslexia but only influence the time at which a dyslexic may be identified. For example, birth trauma may result in an inability to read, but this inability is symptomatic of a localized lesion, and in specific dyslexia no lesion is apparent. Similarly, eye-movement patterns exhibited by specific dyslexics are not the cause of their reading problem but are judged to result from the reading problem.

The outlook in this book is said to be neurological. This approach in *The Dyslexic Child* is somewhat disappointing in that it results in dyslexia being defined not in terms of what it is but rather in terms of what it is not. The book serves mainly to present the author's point of view, although alternative positions are usually adequately discussed and if not, are adequately referenced. The psychological factors contributing to reading failure are outside the scope of this work. However, it should be useful to psychologists interested in a contemporary neurological view of reading disability and of cognitive disorders in general.

Dean G. Purcell, *University of Arkansas*

The Waysiders

By R. M. N. Crosby, with Robert A. Liston. New York: Delacorte Press, 1968. Pp. xiv, 224. \$6.00.

The Waysiders contains an elementary discussion of reading that encompasses both the acquisition of reading skills and the failure to acquire them. The book is written primarily for educators and parents, and to this end it is more than adequate.

The authors define the process of reading within what is called a 'neurological' context. The case studies—used to illustrate various factors that may contribute to reading failure—are chosen to complement this approach. Psychological antecedents of reading failure are not considered.

The neurological approach used is somewhat disappointing for the experimental psychologist. It consists of little more than acknowledgment of the obvious fact that the brain is employed in the job of reading. The hypothesized steps in reading are schematized after the fashion of information-processing theory and would perhaps be better viewed in that context. However, this book should be instructive to the lay reader in that it provides ample evidence for the contention that reading is a complex task and that difficulty can occur for numerous reasons and in various stages of reading development.

Dean G. Purcell, *University of Arkansas*

Short-term Changes in Neural Activity and Behaviour

Edited by G. Horn and R. A. Hinde. Cambridge: Cambridge University Press, 1970. Pp. vii, 627. \$28.50.

This book, which is a series of review papers by acknowledged authorities, is the outcome of a conference on "short-term changes at all levels from the

behavioural to the molecular" (p. vii) that was held at King's College in 1970. Although habituation is emphasized, all types of decremental and incremental behavioral changes are considered in detail. There is so much material in the book that even experts in the area will find much that is new and helpful. The book will be particularly useful as a source of new material and references for those interested in neurophysiology, physiological psychology, or the neural correlates of animal behavior.

Hinde sets the tone in the first chapter, in which he reviews habituation at the behavioral level. Several such reviews have recently appeared in the psychological literature, but they have tended to focus on habituation in higher vertebrates, especially mammals (see reviews by R. F. Thompson, for example). Just as many of the concepts of neurophysiology that were developed in the 1950s have been challenged by the recent work on a variety of creatures (especially molluscs and arthropods), so have the traditional ideas of habituation been challenged by Hinde and others in this book. The more recently acquired concepts in neurophysiology occurred as the natural consequence of an ever-expanding technology; the information on habituation, on the other hand, has always been available, but it is only very recently that it has been recognized by most psychologists. Hinde (and others in the book) points out major exceptions to each of the nine parametric characteristics of habituation generally recognized by psychologists and suggests that these characteristics should be little more than a starting point for any meaningful discussion of habituation (p. 4).

In the subsequent chapters, there is a wealth of material covering all forms of plasticity of behavior, focusing on the neuronal or molecular level. There is almost too much material for one to digest, even given the great amount of redundancy in the book. Several of the chapters are little more than restatements of their authors' earlier papers (those by Weiskrantz, by Bureš and Burešová, and by Griffith, for example). Many more are excellent reviews that tend to focus on the research of the authors and which are all well-supplemented with recent references. These chapters have an excellent overall level of quality, both in written style and in pointing out the goals and the relevance of the research (see those by Vinogradova, Rowell, by Kandel et al., and Rose).

The book includes chapters on habituation, learning and extinction, post-tetanic potentiation and depression, heterosynaptic facilitation, and DC shifts. For all of these topics, there is good collation of behavioral, neuro-anatomical, and neurophysiological data from both vertebrates and invertebrates. Several more theoretically oriented chapters are also included. In the epilogue, Horn concludes that little is presently understood of the mechanisms underlying any of these changes in behavior, with the possible exception of habituation. Yet the reader is left with the impression that such understanding will be forthcoming on many fronts.

The book would be easier to follow if the authors had been required to use a standard terminology; but this drawback, along with the redundancy of most of the topics, is overcome by the overall refreshing approach and review of this frontier area of neuropsychology.

Stephen C. Woods, *Columbia University*

Research and Report Writing in the Behavioral Sciences

By Robert L. Noland. Springfield, Ill.: Charles C Thomas, 1970. Pp. vii, 98.

This unusual little volume should be useful to any undergraduate major in the behavioral sciences. Our reasons are several. Some universities do not educate freshmen in the resources of the library and in how books are given entries in the card catalog. This book does. Guides for research in the library are ample, as are references to scientific method, designs for research in general, survey studies, developmental studies, case studies, clinical studies, and sources of information about research and writing. The principal strength of this book is its very nice collection of references that should save students innumerable hours in their libraries.

Olin W. Smith and Patricia C. Smith, *Bowling Green State University*

Readings in Child Development and Personality

Edited by Paul H. Mussen, John J. Conger, and Jerome Kagan. 2d ed. New York: Harper & Row, 1970. Pp. xii, 595. \$5.95.

The present edition is an updated and expanded version—only four selections remain from the first edition, which contained 115 fewer pages—of a book of readings that was designed as a softcover companion volume to the latest (third) edition of the editors' *Child Development and Personality*, one of the most widely used textbooks in child-psychology courses today.

The revised collection "was designed to make readily available a set of essays and empirical papers, most of them written recently, representing current theory and knowledge in the field of child development" (p. ix). The 34 readings are organized topically and also are arranged into parts designed to illustrate the child's successive stages of development: (I) "Biological and Environmental Factors in Development in the First Two Years," (II) "Early Language and Cognitive Development," (III) "Familial Influences on Socialization and Personality Development," (IV) "Extrafamilial Influences on Personality Development," (V) "The School as an Agent of Socialization," and (VI) "Adolescence." Brief introductions to the sections, with editors' comments about the papers included in each, and separate author and subject indexes greatly add to the book's usefulness.

In addition to the editors, who either authored or coauthored five of the selections, contributors to the present edition include, among others, Bandura, Bellugi, Brown, Erikson, Gottesman, Gewirtz, the Harlows, Hetherington, Moss, Piaget, and White. While the assembling of any book of readings is always a matter of personal preference, several eminent psychologists working in the developmental-personality area are conspicuous in the present collection only by their absence. Nevertheless, most of the selections included in the present edition are excellent; for this reason, the book can be recommended as supplementary reading for almost any course in human development—even for those which do not use its companion volume as a textbook.

Jon D. Swartz, *University of Texas at Austin*

Personality

By Richard S. Lazarus. 2d ed. Foundations of Modern Psychology series. Englewood Cliffs, N. J.: Prentice-Hall, 1971. Pp. 186. \$5.75.

This brief text is designed for introductory courses in psychology and could very effectively be used with other materials to replace the encyclopedic text or to supplement those texts giving superficial treatment to the study of personality theory. Although the attempt is to replace the breadth of the traditional text with depth, the volume is not greatly successful in doing so. Rather, it primarily expands the breadth of coverage. Within less than 200 pages, Lazarus covers basic assumptions of the field; psychoanalytic, neo-analytic, Piagetian, Lewinian, and learning theory approaches; the problem of motivation; biological antecedents; social-influence pressures; assessment; and more. The coverage in each area is so brief that the inquiring student will feel frustration at being offered a heady brew without being allowed to drink. In keeping with traditional introductory treatments, a great emphasis is placed on defining concepts and describing a representative experiment. The good student will thus absorb a great many bits of information but will experience little excitement at intellectual discovery.

The volume is refreshing in its emphasis on experimental research and is written with both clarity and brevity of style. The chapter on biological bases is a welcome one in a field too much dominated in recent years by the environmentalist approach. My applause would be louder had this section been expanded at the expense of the Piagetian theory and other less essential material. The absence of material on self-concepts and on situational shifts in personal dispositions are unfortunate lacunae in the coverage. The volume does touch on issues of broad social relevance. Such materials may compensate for the difficulty many students will have personally identifying with major lines of argument.

Kenneth J. Gergen, *Swarthmore College*

A Cornell Reader

By Raymond F. Howes. Ithaca, N. Y.: Cornell Alumni Association, 1971. Pp. xv, 201. \$5.00.

Although this little book as a whole will be of most direct interest to people who have been students or served on the staff at Cornell University, the section dealing with E. B. Titchener will interest all psychologists. Mr. Howes was the son of Titchener's sister, was born in Ithaca, and after graduating from Cornell in 1924, returned as a graduate student and then held a number of administrative posts at the university over a period of years. His basic interests were in the fields of English and rhetoric, and his discussions of these departments and the men teaching in them at Cornell are particularly full and informative, but his reminiscences of the personality and the social behavior of his famous uncle make very entertaining reading.

Some of the material about Titchener will be known to readers of Boring's *History*: his habit of wearing his Oxford Master's gown when giving the lectures in Psychology 1, for example. But the story about his refusal to be in-

roduced to an ex-student of G. Stanley Hall, a student who had reneged on a promise to complete his dissertation, is not widely known. Titchener said, "I will not speak to such a person." There are several examples of his intellectual integrity, as well as of his rigid code of etiquette as an English gentleman.

Since Mr. Howes is not a psychologist, he makes no attempt to discuss Titchener's contributions to the field nor to evaluate him as a scientist, but his stories give one considerable insight into some of the reasons for Titchener's high standards for scientific work as well as for his rigidity.

Dorothea E. Johannsen, *Tufts University*

Books received

(The books listed here have not as yet been noted in our pages. Their listing does not preclude their later review, however.)

- Abse, D. Wilfred. *Speech and Reason: Language Disorder in Mental Disease*. Charlottesville: University Press of Virginia, 1971. Pp. xiii, 310. \$12.00.
- Abt, Lawrence Edwin, and Riess, Bernard F., eds. *Clinical Psychology in Industrial Organization*. New York: Grune & Stratton, 1971. Pp. viii, 193. \$12.50.
- Akhurst, B. A. *Assessing Intellectual Ability*. New York: Barnes & Noble, 1971. Pp. ix, 180. \$2.00.
- Ancell, Melvin. *Sex and Sanity*. New York: Macmillan, 1971. Pp. viii, 310. \$6.95.
- Bartholomew, D. J., and Morris, B. R., eds. *Aspects of Manpower Planning*. New York: American Elsevier, 1971. Pp. x, 130. \$7.50.
- Blanton, Smiley. *Diary of My Analysis with Sigmund Freud*. New York: Hawthorn Books, 1971. Pp. 141. \$7.95.
- Block, Jack. *Lives through Time*. Berkeley, Calif.: Bancroft Books, 1971. Pp. xxii, 313. \$12.50.
- Brakel, Samuel J., and Rock, Ronald S., eds. *The Mentally Disabled and the Law*. Rev. ed. Chicago: University of Chicago Press, 1971. Pp. xix, 487. \$12.95.
- Buckhout, Robert, ed. *Toward Social Change: A Handbook for Those Who Will*. New York: Harper & Row, 1971. Pp. xiv, 480.
- Buhler, Charlotte. *The Way to Fulfillment: Psychological Techniques*. New York: Hawthorn Books, 1971. Pp. xviii, 237. \$8.95.
- Burroughs, G. E. R. *Design and Analysis in Educational Research*. Educational Monographs, 8. Birmingham, Ala.: University of Birmingham School of Education, 1971. Pp. xii, 327.
- Cancro, Robert, ed. *The Schizophrenic: An Annual Review*. New York: Brunner/Mazel, 1971. Pp. xv, 791. \$20.00.

continued on page 188

ANNOUNCEMENT

Contemporary Views of Learning and Conditioning

Under the chairmanship of F. J. McGuigan (Hollins College) and D. Barry Lumsden, a conference on Contemporary Views of Learning and Conditioning was held May 23-26, 1972, at North Carolina State University, Raleigh. Among the scheduled papers were

William K. Estes, Rockefeller University
"Memory and Conditioning"

W. Horsley Gantt, Pavlovian Research Laboratory, V.A. Hospital, Perry Point, Maryland
"Does Teleology Have a Place in Conditioning?"

David Grant, University of Wisconsin
"Reification and Reality in Conditioning Paradigms: Implications of Results When Modes of Reinforcement Are Changed"

Gregory A. Kimble, University of Colorado
"Scientific Psychology in Transition"

F. J. McGuigan, Hollins College
"Conditioning of Covert Behavior: Some Problems and Some Hopes"

David Premack, University of California at Santa Barbara
"Cognitive Principles?"

Robert Rescorla, Yale University
"On Conditioning"

Eli Saltz, Wayne State University
"There Never Has Been a Learning Theory"

William N. Schoenfeld, Queens College
"The Learner and the Learned" (with Brett Cole, John Lang, and Robert Mankoff)

Martin E. P. Seligman, University of Pennsylvania
"Cognition and Learning Theory" (with James Johnston)

Richard L. Solomon, University of Pennsylvania
"Opponent-Processes of Motivation in Relation to Conditioning"

Delos D. Wickens, Ohio State University
"Classical Conditioning as a Tool for the Investigation of Processes in Various Psychological Areas"

For further information write Box 5096, Raleigh, North Carolina 27607
Sponsored by The Adult Learning Center, D. Barry Lumsden, Director

NON-VERBAL COMMUNICATION

Edited by ROBERT A. HINDE

A collection of edited papers, with linking commentaries and notes, from the work of a Royal Society Study Group set up under the direction of W. H. Thorpe to investigate non-verbal communication and its importance in psychology, zoology, ethology, sociology and anthropology. Plates, diagrams, tables. \$17.50

ETHOLOGICAL STUDIES OF CHILD BEHAVIOR

Edited by N. BLURTON JONES

Studies by research workers in zoology, psychology and psychiatry on social interactions between child and child and between mother and child. The papers describe recent investigations and some still in progress. Diagrams, tables. \$19.50



Cambridge University Press

32 East 57th Street, New York, N.Y. 10022

ANNOUNCEMENT

New Citation Index for Social Sciences

A new index designed to improve access to information published in the social sciences and humanities journal literature has been announced by the Institute for Scientific Information (ISI) in April 1972.

Beginning in 1973, ISI will offer a *Social Sciences and Humanities Citation Index* covering articles, book reviews, and correspondence appearing in approximately 600 key journals in the fields of anthropology, community health, demography, economics, educational research, ethnic studies, geography, history, law, linguistics, management, marketing, political science, psychiatry, psychology, sociology, statistics, and urban planning/development.

Plans for the new index were disclosed by Melvin Weinstock, ISI Senior Information Scientist, at the recent conference on Access to Knowledge and Information in the Social Sciences and Humanities, sponsored by Queens College of the City University of New York. Weinstock said that the social sciences index will be patterned after ISI's widely used *Science Citation Index*®. He claimed that the new reference tool is the result of continuing requests from social scientists to apply citation indexing to the literature relevant to their work.

Citation indexing is based on the fact that current articles referencing the same earlier publication have a subject relationship not only with the earlier work but also with each other. To utilize this relationship, a citation index lists together all articles published during a specified time period that reference the same earlier publication. In effect, the earlier work becomes an indexing term by which someone who is familiar with its content may look up new articles on the same subject.

Weinstock said that "terminology in the social sciences and humanities is particularly imprecise and subject to rapid change. This makes it very difficult for traditional 'subject-heading' indexing systems to draw together related material. Citation indexing, however, lets the best qualified person — the author — index his own work through his reference citations."

Quoting from a study by Maurice B. Line, Director of England's National Central Library, Weinstock told the conference that the "majority of social scientists questioned said they would find a citation index very useful in their literature searching." Weinstock also said that ISI has been conducting its own research on the need for the new citation index and what characteristics it should have in terms of format, coverage, frequency of issuance, cost, etc. He said that the final configuration of the index will reflect the results of this research.

Anyone wishing more information or wanting to make suggestions on any aspect of the *Social Sciences and Humanities Citation Index* should call or write Melvin Weinstock, Institute for Scientific Information, 325 Chestnut Street, Philadelphia, Pennsylvania 19106.

*Selected as an Outstanding Education Book
by Education Book List*

Psycholinguistic Learning Disabilities: Diagnosis and Remediation

BY SAMUEL A. KIRK AND WINIFRED D. KIRK

THIS book will assist those using the *Illinois Test of Psycholinguistic Abilities* as they interpret test results and organize remedial programs.

The authors first discuss the broad area of learning disabilities and explain the development and meaning of the subtests of the *ITPA*. In a presentation of selected studies conducted with the *ITPA*, diagnostic-remedial teachers and psychologists will find research results of value in diagnosis and interpretation of subtest scores and patterns of disabilities. A number of studies show which subtests of the *ITPA* are correlated with reading disabilities and which are not. Other studies indicate particular deficits or assets which are found in different ethnic groups, the mentally retarded, the cerebral palsied, and other children. Three chapters deal with the diagnostic process, both in general and specifically with the analysis and interpretation of *ITPA* scores. The authors describe in detail sample patterns of disabilities and the procedures by which these patterns are developed. "Suggestions for Further Analysis and Diagnosis" alerts the diagnostician to broader aspects of behavior. The last two chapters offer general guidelines for remediation and specific diagnosis of disabilities.

SAMUEL A. KIRK is professor of special education, University of Arizona, and professor emeritus, University of Illinois. WINIFRED D. KIRK is a psychoeducational diagnostic-remedial specialist.

Cloth, \$7.95; paper, \$2.45

UNIVERSITY OF ILLINOIS PRESS *Urbana, Chicago, London*

INFORMATION THEORY AND ESTHETIC PERCEPTION

by Abraham Moles, translated by Joel E. Cohen. Here is the first English translation of Moles's *Théorie de l'information et perception esthétique*, a modern pioneering work that greatly enlarges the role of information theory by applying it to perception, particularly esthetics and the psychology of perception. "There is boldness and imagination in this book." — *Science*. "Prof. Moles merits high regard for his important contributions to our understanding of aesthetic experience by emphasizing the materiality of communication . . ." — *Philosophy and Phenomenological Research*. 1966. 217 pp. Cloth, \$7.50; paper, \$2.45.

THE MEASUREMENT OF MEANING

by Charles E. Osgood, George J. Suci, and Percy H. Tannenbaum. The authors deal with the nature and theory of meaning and present a new, objective method for its measurement called the *semantic differential* — a general technique of measurement that can be adapted to a variety of problems and thus has far-reaching implications for empirical research. "Drawing liberally from the researches of clinician and experimentalist alike, Osgood has formulated a problem area which may well provide the much needed basis upon which the integration of psychology can be effected." — *The Canadian Journal of Psychology*. 1967. 342 pp. Paper only, \$3.45.

OBJECTIVE PERSONALITY AND MOTIVATION TESTS

A Theoretical Introduction and Practical Compendium

by Raymond B. Cattell and Frank W. Warburton. Part I treats the theoretical basis and nature of objective personality tests, including historical data, an indexing system for tests and behavioral variables, psychometric theory, test taxonomy, and principles of test design. Part II contains 612 personality tests, each indexed and described to provide maximum information on its design and usage. ". . . highly informative and provocative . . . a valuable source of material." — *American Journal of Psychiatry*. 1967. 687 pp. \$20.00.

PROPOSED MODEL FOR VISUAL INFORMATION PROCESSING IN THE HUMAN BRAIN

by Matthew Kabrisky. An attempt to decipher some of the operation and calculation techniques of the human brain. The book first describes the nervous system of the vertebrate animal analyzed as an information processing device with particular emphasis on the human visual system. It is then shown that known biological structures could be capable of performing such a calculation as a two-dimensional cross correlation. ". . . makes a valuable contribution." — *Quarterly Review of Biology*. 1966. 56 figures. 104 pp. \$6.75.

University of Illinois Press Urbana Chicago London

"What forces us to believe that our only choice is between war or surrender? What is there about human thinking under stress that pushes every group conflict toward mutual annihilation? . . . Let us look into the dynamics of our own minds."

AN ALTERNATIVE TO WAR OR SURRENDER

by Charles E. Osgood

Here is a brilliantly conceived, carefully blueprinted plan for an approach to international relations in the nuclear age that is as fresh now as it was when the book was first printed in 1962 because it is psychologically sound. It offers a way out of the dangerous dilemma posed by the increasing tensions of the arms race that is leading to a permanently divided world, if not a fatal nuclear war.

This plan, which the author has named *Graduated Reciprocation in Tension-reduction* — GRIT — assumes that the problems we face are primarily matters of human nature and human relations and that the Soviet people and their leaders are more like us than the "bogey man" our psychology creates.

It provides that the United States take the initiative in reversing the arms race and in reducing world tensions through a carefully planned and executed "peace offensive" designed to invite reciprocation from the Communist bloc. It is a step-by-step approach with built-in safeguards aimed at maintaining our security.

"Let us look into the dynamics of our own minds, because only through understanding ourselves can we ever hope to understand others." IB-7. 183 pages. Paper, \$1.45.



UNIVERSITY OF ILLINOIS PRESS
Urbana Chicago London

NEW PHOTO-ELECTRIC PURSUIT ROTOR



The new Pursuit Rotor is manufactured in two models, the MK3 and the MK3/T which incorporate a digital 'time on target'. The variable speed range is from 5 to 50 revolutions per minute and the sampling periods are standardised on 5 — 10 and 20 seconds for simplicity of operation. The MK3/T, as depicted, has the above features and also includes a 4 digit totalising counter to eliminate the use of an external 'total time on target' counter.

SINGLE CHANNEL GROUP TACHISTOSCOPE



The instrument is a general purpose slide projector with remote control facilities and interchangeable lenses. The projector has been modified by the addition of a high-speed galvanometer which operates a shutter membrane to obscure the light path. The rise time of the shutter is 1.5 milli-seconds. This instrument can be used in word, picture and symbol recognition and subliminal research.



Forth Instruments Ltd.,
Engine Road, Loanhead, Midlothian,
Scotland.
Telephone: Loanhead 104.

Manufacturers of Single Channel Group Tachistoscopes, Pulse Generators, Memory Drums, Counter Timers, Voice Switches, Micro Electrode Pullers, Reaction Testers.

APC
Dorothy R. Williams
194ue
Librarian
27/4/73

THE AMERICAN

JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys
William E. Kappauf

M. E. Bitterman
Benton J. Underwood

-
- WILLIAM P. WALLACE. Estimations of Event Frequency and the Re-pairing Decrement 311
- G. J. JOHNSON. Sequential and Positional Cues in Serial to Paired-Associate Transfer 325
- CYNTHIA TURNURE. Perceptual Learning in Young Children: Varied Context and Stimulus Labels 339
- MICHAEL STADLER. Figural Aftereffects as Optical Illusions? 351
- LOUIS BUFFARDI. Factors Affecting Memory Span in Binary and Octal Responding 377
- RUTH F. DEICH. Color Recognition and Word Frequency 393
- E. SCOTT GELLER and CHARLES P. WHITMAN. Probability Learning When Stimuli Vary on Multiple Dimensions 397

continued on back cover

Editors Lloyd G. Humphreys M. E. Bitterman
William E. Kappauf Benton J. Underwood

Book Review Editor Julian Hochberg

Consulting Editors William Epstein Robert Fox Wendell Garner
Lawrence E. Jones John Krauskopf Michael Rashotte Allan Wagner

Index Editor Rand B. Evans Necrology Editor E. R. Hilgard

The *American Journal of Psychology* was founded in the interest of general experimental psychology and publishes reports of original research, shorter notes and discussions, and pertinent book reviews. Manuscripts of articles and notes may be submitted to Drs. Humphreys or Kappauf (University of Illinois at Urbana-Champaign, Champaign, Illinois 61820), Bitterman (University of Hawaii, Honolulu, Hawaii 96822), or Underwood (Northwestern University, Evanston, Illinois 60201). Books intended for review should be sent to Dr. Carl P. Duncan (Northwestern University, Evanston, Illinois 60201), who will be the new Book Review Editor as of 1973.

Authors of articles may order offprints at the time they return galley proof. Other contributors are sent tear-sheet copies as available unless extras are ordered when manuscript is accepted.

Contributors are asked to follow the style described in the *Publication Manual* of the American Psychological Association. Keep abstract and title quite brief, use footnotes only when essential, avoid all but the most common abbreviations (2 mm, 4 deg, 60 db, 15 sec; *but* subjects, experimenter, paired-associate learning), and use the tables and figures to supplement rather than repeat the text or each other; a well-designed table or figure can provide more information than many words of text. If you elect to omit from your manuscript any information that is critical to the evaluation of your data (analyses of variance summary tables, means, variances, correlation and factor matrices), provide two copies of that material for the editors' and their consultants' use in reviewing the paper.

The subscription price is \$10.00 a year (foreign \$10.50); single copies cost \$2.75. Back numbers and volumes may be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which all business inquiries (subscriptions, advertisements, changes of address) should be addressed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

© 1972 by the Board of Trustees of the University of Illinois
Manufactured in the United States of America

21 JUN 1973

LAWRENCE E. MARKS and JOSEPH C. STEVENS. Perceived Cold and Skin Temperature as Functions of Stimulation Level and Duration	407
K. G. WHITE. Implicit Contours in the Zöllner Illusion	421
EUGENE B. ZECHMEISTER. Orthographic Distinctiveness as a Variable in Word Recognition	425
FRANK DAPOLITO, DOUGLAS BARKER, and JOSEPH WIAIT. The Effects of Contextual Changes on Component Recognition	431
Book Reviews (Adam Kendon, Paul C. Vitz, Stanley Coren, Kenneth H. Kurtz, Kenneth J. Gergen; books received)	441

With Vol. 86, No. 1 (March 1973)
The American Journal of Psychology
will increase its pages
from 160 per issue (640 a year)
to 224 per issue (896 a year)

AAMD

ADAPTIVE BEHAVIOR SCALES

THE ABS is an Adaptive Behavior Scale for mentally retarded and emotionally maladjusted individuals. It rates the effectiveness of the individual in coping with the natural and social demands of his environment. THE ABS MEASURES individual skills and habits including independent functioning, physical functioning, language development, number and time concepts, and others; maladaptive behavior related to personality and behavior disorders as in violent and destructive behavior, withdrawal, unacceptable or eccentric habits, self-abusive behavior, sexually aberrant behavior, and others. DEVELOPED under a grant from the National Institute of Mental Health, the Scales are now in use in all of the 50 states and many foreign countries. A SPECIMEN SET is available which includes Scales for both Children and Adults, a Manual of Instructions, Answer Sheets, and Hand Scoring Sheets. Price — \$7.50. Please prepay. ORDER FROM THE AMERICAN ASSOCIATION ON MENTAL DEFICIENCY, 5201 Connecticut Avenue, N.W., Washington, D.C., 20015.

ESTIMATIONS OF EVENT FREQUENCY AND THE RE-PAIRING DECREMENT

William P. Wallace

University of Nevada, Reno

A frequency-estimation task intervened between acquisition of a verbal-discrimination list and a delayed test trial on the same pairs or a re-paired list. There was a greater increase in errors on the re-paired list. The frequency judgments showed that re-pairing both increased the number of test pairs for which the wrong had a higher frequency than the right and produced a greater absolute difference in frequency between the rights and wrongs.

A recent analysis of verbal discrimination led to the proposal of some minor modifications in the frequency theory of verbal-discrimination learning (Wallace, 1972). The present experiment was concerned with the validity of the assumptions involved in that proposal.

Verbal-discrimination tasks generally involve the presentation of a series of pairs of items, with one item in each pair designated as correct. Acquisition is said to be complete when subjects can identify the correct item for each pair. According to frequency theory, the subjects accomplish this by judging the difference in frequency between the two members of each pair (Ekstrand, Wallace, and Underwood, 1966). It is assumed that frequency accrues more rapidly to correct alternatives than to incorrect alternatives; hence during the course of learning, the 'rights' become higher in frequency than the 'wrongs.' The theory asserts that subjects discriminate differences in frequency and generally respond by selecting the pair member that has the higher frequency value.

Kanak and Dean (1969) introduced a modification in the usual procedure. At one point in acquisition they re-paired the correct and incorrect alternatives, while leaving the respective correct and incorrect functions unchanged. They found that the re-pairing produced a marked impairment in performance and interpreted this as discrepant with frequency theory, since the re-pairing manipulation did not change the correct and incorrect functions of items. Instead, Kanak and Dean argued, their data were consistent with the position that associations between corresponding correct and incorrect alternatives play an im-

portant role in the acquisition of a verbal discrimination (see 1969).

It has been argued elsewhere that the re-pairing decrement is not discrepant with frequency theory (Wallace, 1972; Wallace and Nappe, 1971). However, these arguments are derived from the 'adjusted' version of frequency theory. The major concern of the present research was with the validity of the assumptions on which these arguments were based. It is to these assumptions that this discussion now turns.

The major assumptions critical to accommodating the re-pairing decrement within frequency theory are four. First, it is assumed that there is variability in frequency values within the class of correct alternatives and within the class of incorrect alternatives. That is, even though two (or more) correct alternatives may have the same number of exposures during learning, they may still differ in frequency value.

The second assumption is that learning difficulty increases with the extent to which the correct alternatives of *all* the pairs in a given list do not exceed their incorrect alternatives in frequency value. Special cases such as transfer situations in which the wrongs are all higher than the rights in frequency value will not be considered here.

Third, it is assumed that specific right and wrong items exposed together will be 'yoked' in frequency value. For example, if some extraneous event occurs at the time a given pair is presented, and this event serves to isolate the pair, then the frequency values for both members of this pair will have the isolation advantage.

Finally, since it is still assumed that frequency builds up faster to rights than to wrongs, the rights will during the course of learning become higher than their respective wrongs in frequency value. However, re-pairing may produce some new pairs that combine a relatively low-frequency right with a relatively high-frequency wrong—some pairs for which the wrongs exceed their respective rights in frequency value. Thus, it is less likely that for all pairs the correct alternatives will surpass their corresponding incorrect alternatives in frequency value after re-pairing, and re-pairing should result in verbal-discrimination performance inferior to that when the same pairs are maintained.

The approach to testing these assumptions may now be briefly described. Verbal-discrimination performance on a re-paired list was compared with performance when the original pairs were retained. However, before this test, the subjects were given an absolute-frequency judgment task similar to the one employed by Underwood, Zimmerman, and Freund (1971). The major concern was to examine the relations between the frequency estimates and performance in

verbal discrimination, and to determine if judged frequency conformed to expectations from the four assumptions. The general plan of the experiment involved a verbal-discrimination acquisition phase, a frequency-estimation task, and a final verbal-discrimination test with the acquisition pairs or with items re-paired.

METHOD

The experiment involved three phases, and every subject participated in all three phases. The major variables were frequency of pair exposure (within subjects) in phase 1, frequency-estimation instructions (between subjects) before phase 2, and a re-pairing versus same-pair verbal-discrimination test (between subjects) in phase 3.

—Phase 1—The first phase of the experiment involved one study trial and one test trial on a 33-pair list of common two-syllable English words. The lists were presented on a tape recorder at approximately 5 sec per pair. During the study trial, a female voice on the tape recorder read both members of a pair and then repeated the correct word: for example, "obtain, reduce, obtain is correct." During the test trial, designation of the correct word was omitted.

Two lists were constructed by randomly pairing 130 two-syllable words. In addition, each pair appeared in one list with one member correct and in another list with the other member correct. Thus, four separate tapes were prepared so that two word lists were used and that each member of a pair served as the correct member in one list and as the incorrect member in another list.

Within each list, five different types of pairs may be distinguished. The basis for distinction is the number of times the pairs were read during the study trial. The assignment of specific pairs to these five classes was random. Eight pairs were read *once* each during the study trial, eight pairs were read *three times* each, eight pairs were read *five times* each, and eight pairs were read *eight times* each. The fifth class contained one pair that was the *same* on all four tapes; it was read *four times* during the study trial and provided the 'anchor' word for phase 2 of the experiment. The specific pairs were assigned at random to study-trial and test-trial positions, with the restriction that at least three positions separate the repetition of any pair during the study trial.

The test trial was separated from the study trial by approximately 10 sec, during which the subjects were informed that they would now have a test. During the test the voice on the tape recorder read both members of a pair, and the subjects responded by circling either a 1 or a 2 on their answer sheets (at the appropriate place), depending on whether they wished to select the first or second word read. A random-assignment procedure was used to determine whether the correct alternative was read first or second within a pair.

—Phase 2—The second phase of the experiment followed immediately

after phase 1. The subjects were given a ten-page booklet that had an instruction sheet followed by nine pages, each of the latter with 9 different words arranged in a single column. Each booklet included the 64 words that appeared one, three, five, and eight times during the phase 1 study trial. In addition, there were 16 two-syllable common English words that had not appeared in phase 1. Finally, the first word presented in each booklet was the correct item from the pair presented four times during phase 1 (the anchor word). With the exception of the anchor word, the remaining words were arranged randomly, with two random orders for each of the two lists.

The subject's task during this phase was to assign to each word a number representing his estimate of the frequency with which that word had occurred during phase 1. A magnitude-estimation procedure was employed. The subjects worked at their own pace, and the duration of phase 2 was approximately 10 min.

There was one variation in instruction introduced at this point. Half of the subjects received instructions stressing that they were to base their estimates on the *actual* (A) presentations of words during phase 1. The other subjects were further told: "In addition, you may have spent unequal amounts of time going over or thinking about individual words as they were presented. Your task in this part of the experiment is to estimate the relative frequencies of occurrence, based upon actual presentations and time spent thinking about or rehearsing each word during the first part of the experiment." This instructional variation was included to explore estimations of *subjective* (S) frequencies. In both cases the instructions went on to explain the magnitude-estimation procedure.

The anchor word was included to provide a common baseline for the estimations to the other 80 words used in phase 2. The same operation that was necessary to transform a given subject's estimate on the anchor word to 10 was applied to his estimates for his other 80 words (see Stevens and Tulving, 1957). Thus, every frequency-estimation protocol began with the anchor word and a score of 10.

—Phase 3—The third phase of the experiment involved a verbal-discrimination test trial with the list learned in phase 1. For half of the subjects, the test involved the *same pairs* (groups A-SP and S-SP for the two phase 2 instruction conditions) that were used in phase 1; for the remaining subjects, the test involved *cross pairs* (groups A-CP and S-CP), or re-paired rights and wrongs.

The nature of the cross pairs can be shown by tracing the re-pairing of the eight pairs that were presented eight times each during the phase 1 study trial. Two of these pairs had their correct alternatives ('rights') re-paired with the incorrect alternatives ('wrong') from pairs exposed *one* time each during the phase 1 study trial, two more had their rights re-paired with wrongs from pairs exposed *three* times each, two more had their rights re-paired with wrongs from pairs exposed *five* times each, and the last two had their rights re-paired with wrongs from two other pairs exposed *eight* times each. Pairs that were presented one, three, and five times each were similarly re-paired.

A specific example will illustrate the semirandom procedure used to construct the new, cross pairs. Two pairs each presented eight times during phase 1 (attempt-upon and equal-extend) were selected at random to be re-paired with the words from two pairs each exposed once during phase 1 (marry-remain and flower-glory). Assume that in each of the above pairs the first word is the correct alternative. The new pairs with eight exposures of the correct alternative and one exposure of the incorrect alternative were attempt-remain and equal-glory. The new pairs with one exposure of the correct alternative and eight exposures of the incorrect alternative were marry-extend and flower-upon. The re-pairings were determined in such a way that if the right word from pair A was paired with the wrong word from pair B, then the wrong word from pair A was not paired with the right word from pair B. Of course, new pairs with equal exposure frequencies provided an exception to this rule, since the re-paired items with eight exposures of both rights and wrongs were metal-degree and reveal-shoulder, whereas the original pairs were reveal-degree and metal-shoulder.

During the phase 3 test, the 32 pairs (the anchor pair was not included) were presented in a single column on a sheet of paper. The order of the pairs was random. The pairs were covered by a template. Upon a signal from the experimenter, the subjects slid the template down to expose the first pair and circled the member of the pair they thought was correct. The subjects were instructed to keep up with the experimenter's pace, to expose only one pair at a time, to make a selection for each pair, and not to change earlier responses. The subjects responded at about 3-sec intervals, as signaled by the experimenter.

—Subjects—The subjects were run in groups of from three to eight, with each combination of phase 2 instructions and phase 3 test pairings represented once for every four subjects. There were six subjects in each of the 16 groups that resulted from the combination of the four phase 1 tapes, the two phase 2 instructions, and the two phase 3 test pairings.

The original plan of the experiment called for eight subjects in each of these groups (for a total of 128 rather than 96 subjects), but some subjects were lost due to experimental error and apparent misunderstanding of the phase 2 instructions. Specifically, the phase 2 booklets for three subjects had pages missing, and one subject skipped a page during phase 2. These subjects were discarded because they contributed incomplete data sheets. Ten other subjects were discarded because their frequency rating indicated an apparent failure to understand the task requirements. The most frequent 'failure to understand' was reflected by the use of only two values for the 81 frequency estimates. The instructions had included, as an example of assigning numbers to words, the request that "if you judge the word X to be twice as high in frequency value as the word Y, then the number you assign to X should be twice as large as the number you assign to Y (e.g., $X = 2$, $Y = 1$, . . .)," and several subjects used only 1's and 2's in responding; indeed, one subject assigned either an X or Y to each of the 81 words. The largest number of subjects discarded from any of the 16 groups was two. In the groups that lost less than two subjects, protocols were discarded at random to equalize group size at six.

RESULTS

Phase 1

A $2 \times 2 \times 2 \times 2 \times 4$ analysis of variance was used to compare the errors on the first test trial. Two of the variables in this analysis—phase 2 instructions and phase 3 test pairings—were dummy variables, since the manipulations were not introduced until after the test trial. The other variables were the two word lists and the two item functions (correct and incorrect) that differentiated the four tapes, plus the within-subject variable of pair frequency (one, three, five, or eight presentations).

The analysis resulted in three significant effects. As expected, errors decreased as the number of presentations increased [$F(3, 240) = 29.30, p < .01$]. In addition, there was a significant main effect for materials, as fewer errors were made with one of the word lists

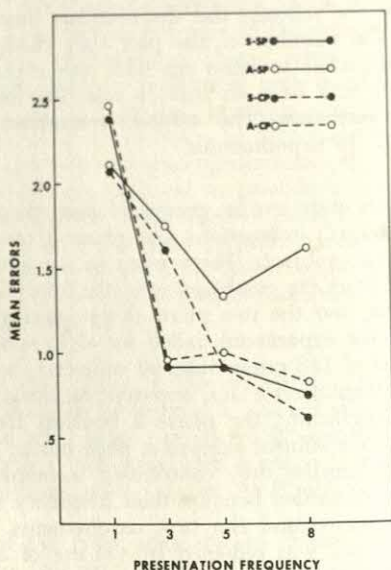


Fig. 1. Mean errors on the phase 1 test trial as a function of presentation frequency, phase 2 instructions (A emphasizing actual, S emphasizing subjective frequency), and phase 3 test pairings (SP for same-pair, CP for cross-pair conditions)

[$F(1, 80) = 7.23, p < .01$]. Finally, there was an unexplained second-order interaction among frequency and the two dummy variables of

phase 2 instructions and the type of test pairing in phase 3 [$F(3, 240) = 3.96, p < .01$].¹ The nature of this interaction can be seen in Figure 1.

Phase 2

One fact evident in the frequency-rating booklets was the considerable variability among the judgments of each subject to words with identical exposure frequencies. This raised two types of questions. The first was how the estimates were influenced by instructions given before phase 2, by presentation frequency in phase 1, and by right and wrong function in phase 1. The second type of question concerned the relation between phase 3 performance and the estimated frequencies of the words that made up the phase 3 test pairs. This latter question will be considered in the next section.

Three preliminary analyses on the frequency estimates preceded the major analysis. The four major groups (phase 2 instructions $\times$ phase 3 test pairings) did not differ on the raw scores that subjects assigned to the anchor word. All subsequent analyses were based upon the transformed estimates, so that if a given subject assigned a value of 5 to the anchor word, each of his remaining estimates was multiplied by 2, the factor necessary to transform the anchor word's value to 10. An $F_{\max}$ test on these transformed scores was not significant, indicating that the assumption of homogeneity of variance was not violated. Finally, the four major groups did not differ on the values assigned to the 16 filler words that were not presented during phase 1. The instructions emphasizing subjective estimates (the S groups) resulted in higher scores, but this effect did not reach acceptable levels of statistical significance [$F(1, 92) = 3.34$].

The effects of four variables were investigated in the major analysis of the transformed frequency judgments. Two of these variables were manipulated between subjects: phase 2 instructions and the dummy variable of phase 3 test pairings. The remaining two variables were manipulated within subjects: phase 1 exposure frequency (one, three, five, or eight presentations) and phase 1 function (right or wrong).

As was true with the filler words, the instructions emphasizing subjective estimates led to higher scores, but again this effect was not significant [$F(1, 92) = 2.59$]. The estimated frequency values increased as a function of phase 1 presentation frequency [$F(3, 276) = 81.63, p < .01$], and the frequency estimates were higher for rights than they were for wrongs [$F(1, 92) = 83.48, p < .01$]. The only other significant effect was for the interaction between phase 1 frequency and

phase 1 function [$F(3, 276) = 8.48, p < .01$]. The nature of this interaction is depicted in Figure 2, which shows that the difference

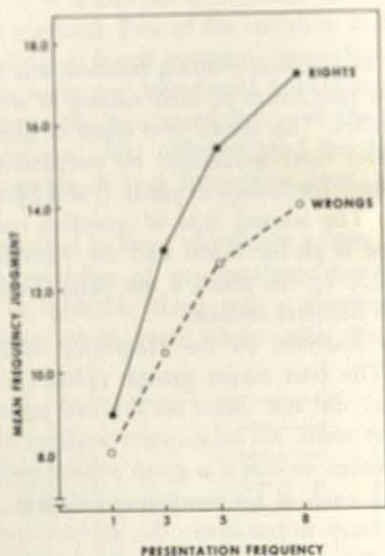


Fig. 2. Mean frequency judgments for correct (R) and incorrect (W) alternatives presented one, three, five, or eight times during phase 1

between right and wrong items in estimated frequency increased as presentation frequency increased. Apparently, number of presentations not only increased base frequency (the frequency value of the less frequent member of a pair) but also increased the absolute frequency difference between the right and wrong alternatives (see Underwood and Freund, 1970).

An attempt was made to evaluate the 'yoking' assumption mentioned in the introduction to this report. It was reasoned that right and wrong items exposed together would be more similar in frequency values than right and wrong items that had not been paired together during the study trial (even though the latter pairs were presented the same number of times as the former pairs). For computational convenience, the present analysis included only the 48 subjects in the conditions with the same pairs (groups SP). Two scores were derived for each subject. The first was the sum of the absolute values of differences in estimated frequency value for the 32 study-trial pairs. The

second score was computed in the same manner on 32 items that had been regrouped so that rights and wrongs with the same presentation frequency were compared together; rights presented eight times were all 're-paired' with wrongs presented eight times, for example. It should be noted that these were *statistical* pairs and had not been presented as pairs in either phase 1 or phase 3.

If the yoking assumption is viable, then the absolute differences in estimated frequency values among these statistical pairs should be lower for the 32 study-trial pairs than for the 32 reconstructed pairs. The average absolute value of the difference in estimated frequency per pair was 6.2 for the study-trial pairs and 6.8 for the reconstructed pairs. Although this difference is relatively small in magnitude, a *t* test for correlated samples indicated it was significant [$t(47) = 3.66$, $p < .01$].

Phase 3

Because of the unexplained interaction in phase 1 of variables that were not manipulated until phases 2 and 3, the major analyses on phase 3 error scores were based on loss scores—on the increase in errors from phase 1 to phase 3, derived by subtracting each subject's total phase 1 errors from his total phase 3 errors. A $2 \times 2 \times 2 \times 2$ analysis of variance compared phase 2 instructions, the same-pair or cross-pair conditions (SP and CP groups) in phase 3, and the word lists and item functions determining the four tapes used in phase 1. Subjects in the same-pair conditions showed a mean increase of 1.46 errors from the first verbal-discrimination test in phase 1 to the second test in phase 3. The mean increase in errors for subjects in the re-pairing conditions was 2.65. The difference between these same- and cross-pair conditions was significant [$F(1, 80) = 4.75$, $p < .05$]. No other effect in this analysis produced an *F* greater than 1.0.

Re-pairing rights and wrongs thus resulted in a greater increase in errors from phase 1 to phase 3; however, a direct comparison of phase 3 errors did not result in a reliable difference. The total errors in phase 1 and phase 3 are shown in Figure 3. The differences between the same- and cross-pair conditions in the slopes of the lines connecting phase 1 and phase 3 errors reveal the greater increase in errors for the cross-pair conditions.

Since number of actual presentations of phase 3 test pairs was ambiguous for the cross-pair groups, two additional analyses were performed to assess the effects of exposure frequency on the phase 3

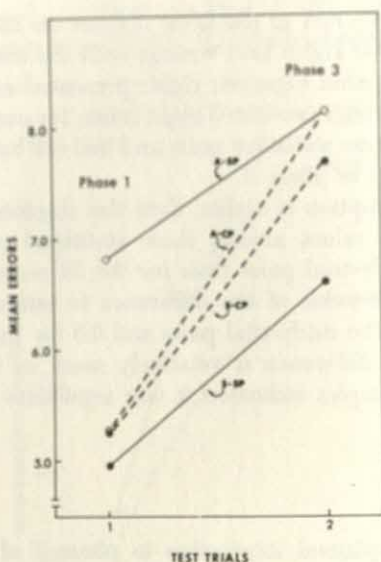


Fig. 3. Mean errors in phase 1 (test trial 1) and phase 3 (test trial 2) as a function of phase 2 instructions (A emphasizing actual, S emphasizing subjective frequency) and phase 3 test pairings (SP for same-pair, CP for cross-pair conditions)

verbal-discrimination test. The first involved only the same-pair conditions (the SP groups). A 2×4 analysis of variance (phase 2 instructions and presentation frequency) indicated a significant main effect for number of presentations during the phase 1 study trial [$F(3, 138) = 16.46, p < .01$]. No other effect was significant.

The second analysis comparing phase 3 errors involved only the cross-pair conditions (the CP groups). Three classifications of pairs, based on number of presentations during phase 1, were created for this analysis: pairs with rights presented more often than their corresponding wrongs ($R > W$), pairs with wrongs presented more often than their corresponding rights ($W > R$), and pairs with equal presentations of rights and wrongs ($R = W$). The phase 3 error scores were transformed to number of errors per opportunity, since there were 12 $R > W$ pairs, 12 $W > R$ pairs, and 8 $R = W$ pairs. A 2×3 analysis of variance (phase 2 instructions and R/W classification) was used to compare the proportion of errors on the phase 3 test. The mean errors per opportunity were .18 for $R > W$ pairs, .22 for $R = W$ pairs, and .33 for $W > R$ pairs. The differences among these conditions were

significant [$F(2, 92) = 14.88, p < .01$]. Again, no other effect was significant.

In the analyses above, errors on the phase 3 test were compared as a function of frequencies of exposure during phase 1. The analyses below involve frequencies of exposure as estimated during phase 2.

It may be recalled that in accounting for the re-pairing decrement, it was assumed that re-pairing rights and wrongs increases the likelihood that some wrongs will have higher frequency values than their corresponding rights. The analysis of loss scores showed that there was a re-pairing decrement. At this point we ask whether re-pairing altered the relations between respective rights and wrongs in estimated frequency value.

Specifically, the first question here is whether re-pairing increased the number of pairs for which the wrong item had a higher estimated frequency than its corresponding right item. To answer this question, each subject's frequency estimates to the members of his phase 3 test pairs were tallied, and the number of pairs for which the frequency he had assigned the wrong word exceeded the frequency he had assigned the right word was determined. The number of such pairs provided the basic data for a 2×2 analysis of variance (phase 2 instructions and phase 3 test pairings). This analysis resulted in one significant effect. There were more pairs for which the wrong word had a higher frequency value than the right word in the cross-pair groups than in the same-pair groups [$X = 10.31$ and 8.60 respectively; $F(1, 92) = 7.93, p < .01$].

Given that cross-pairing did indeed produce more $W > R$ pairs, the next question is whether performance on the phase 3 test trial was related to the difference between rights and wrongs in estimated frequency value. Two scores were derived for each subject. These scores were the proportion of $R > W$ pairs that were correct on the phase 3 test and the proportion of $W > R$ pairs that were correct. The protocols of all subjects had both types of pairs. Neither phase 2 instructions (A or S) nor phase 3 pairing condition (SP or CP) affected these scores. The only significant effect in the analysis resulted from the higher proportion correct for $R > W$ pairs than for $W > R$ pairs [$X = .83$ and $.62$ respectively; $F(1, 92) = 99.56, p < .01$].

A final, related question involves comparing frequency estimates of the rights and wrongs of those phase 3 pairs to which the subjects responded correctly with those to which they responded to by selecting the incorrect alternative. The average difference in estimated frequencies (rights minus wrongs) was computed for pairs responded to

correctly in phase 3 and for pairs responded to incorrectly. One subject made no errors and was assigned a value of zero for incorrect pairs. A $2 \times 2 \times 2$ analysis of variance (the four major groups and correct versus incorrect pairs) revealed that the average frequency difference for pairs responded to correctly was significantly higher than for pairs responded to incorrectly [$X = 6.73$ and -1.50 respectively; $F(1, 92) = 83.30$, $p < .01$]. The only other significant effect indicated that the average frequency difference for the same-pair conditions was greater than the average frequency difference for the cross-pair conditions [$X = 1.81$ and $.81$ respectively; $F(1, 92) = 4.09$, $p < .05$].

DISCUSSION

The present study found a greater increase in errors on a verbal-discrimination test with re-paired rights and wrongs than on one with the original pairs. As in an earlier report (Wallace and Nappe, 1971), the magnitude of the effect was small. In addition to this demonstration of a re-pairing decrement, there was evidence in support of a frequency-theory interpretation of the effect.

It seems reasonable to assume that the states of the environment and the organism are continually varying during the course of verbal-discrimination learning. At any point in time, these conditions may be more or less favorable to acquisition (frequency accrual). Hence, there ought to be certain similarities in frequency values to rights and wrongs of pairs presented together, since they share the same environmental and organismic influences.

This 'yoking' is assumed to result in specific pairs with members relatively close in frequency value. If frequency accrues faster to rights than to wrongs, at some point in learning the rights will all be higher in frequency value than their respective wrongs. Verbal discrimination should be impaired to the extent that all pairs do not conform to the rule that the correct alternative is higher in frequency than its corresponding incorrect partner. Re-pairing rights and wrongs dissolves the yoking advantage and may result in some new pairs that combine a wrong item of relatively high frequency value with a right item of relatively low frequency value. Such re-pairing reduces the likelihood that all the rights are higher than their respective wrongs, and consequently performance is impaired.

Implicit in the discussion above is the assumption that frequency values are variable among items with similar right or wrong functions

and identical numbers of presentations. Certainly, the frequency-judgment protocols were consistent with this assumption, although, of course, the variability may be intrinsic to the measuring instrument.

It is apparent that several assumptions are necessary to incorporate the re-pairing decrement within frequency theory. The evidence offered in support of the validity of these assumptions was based upon subjects' estimates of frequency and the relation of these estimates to verbal-discrimination performance. The frequency judgments were sensitive to differences in actual number of presentations of pairs and the right or wrong function of individual words.

The frequency judgments also revealed an interesting interaction between number of presentations and item function: the difference in estimated frequency between rights and wrongs increased as number of presentations increased. Underwood and Freund (1970) reported that the ability to discriminate differences in frequency was poor when base frequency (number of presentations of the less frequent member of a pair) was greater than three, the implication being that frequency discrimination should be difficult beyond three presentations of a pair. The present results may indicate that the greater increase in frequency differences as a function of number of presentations compensates for the difficulty that accompanies increases in base frequency.

There was some evidence that the yoking assumption, which is central to the present frequency-theory interpretation of the re-pairing decrement, is tenable. Randomly re-pairing rights and wrongs (which were exposed the same number of times in phase 1) resulted in a greater difference in the absolute values of estimated frequency between those pair members, as compared to the difference in absolute values for the members of the original pairs. This is consistent with the assumption that exposing words together minimizes the frequency differences between right and wrong words.

The re-pairing decrement was attributed to an assumed increase in the number of pairs for which the rule to select the more frequent member of the pair would result in an error. The frequency estimates provided data consistent with this assumption. There were more test pairs with wrongs greater in frequency value than their respective rights ($W > R$) in the re-paired conditions, compared to the same-pair conditions. Re-pairing did not alter the probabilities of correct responding to $W > R$ pairs or to $R > W$ pairs. The cross-pairing manipulation just increased the number of the more difficult, $W > R$, pairs.

Magnitude estimation of frequency value appeared to be a useful approximation to the assessment of the relations between frequency

and verbal-discrimination learning. Instructional variations presumed to emphasize a subjective basis for frequency judgments had little influence. In view of the variability in frequency accrual to individual words, it is argued here that judged frequency provides a more effective approximation of 'true' frequency values than is provided by the counting postulate introduced by Ekstrand et al. (1966).

Notes

This research was made possible by Grant 1 RO3 MH19323-01 from the United States Department of Health, Education, and Welfare. Received for publication July 1, 1971.

1. This interaction was significant when the performance of the original 128 subjects was compared [$F(3, 336) = 3.42, p < .05$], indicating that the interaction was not due to some unknown factors connected with the attrition of subjects.

References

- Ekstrand, B. R., Wallace, W. P., and Underwood, B. J. 1966. A frequency theory of verbal-discrimination learning. *Psychological Review* 73:566-578.
- Kanak, N. J., and Dean, M. F. 1969. Transfer mechanisms in verbal discrimination. *Journal of Experimental Psychology* 79:300-307.
- Stevens, J. C., and Tulving, E. 1957. Estimations of loudness of a group of untrained observers. *American Journal of Psychology* 70:600-605.
- Underwood, B. J., and Freund, J. S. 1970. Relative frequency judgments and verbal discrimination learning. *Journal of Experimental Psychology* 83:279-285.
- Underwood, B. J., Zimmerman, J., and Freund, J. S. 1971. Retention of frequency information with observations on recognition and recall. *Journal of Experimental Psychology* 87:149-162.
- Wallace, W. P. 1972. Verbal discrimination. In *Human memory: Festschrift for Benton J. Underwood*, ed. C. P. Duncan, L. Sechrest, and A. W. Melton. New York: Appleton-Century-Crofts, in press.
- Wallace, W. P., and Nappe, G. W. 1971. Re-pairing "rights" and "wrongs" in verbal discrimination learning. *Journal of Experimental Psychology* 87:355-360.

SEQUENTIAL AND POSITIONAL CUES IN SERIAL TO PAIRED-ASSOCIATE TRANSFER

G. J. Johnson

University of British Columbia

The relative positive transfer—from a serial list to a paired-associate list—that was mediated by interlist positional relations was almost twice as pronounced as that mediated by interlist sequential relations. When both positional and sequential cues were available, transfer was no greater (except for paired-associate items from the middle of the serial list) than that with positional cues alone.

In an attempt to assess the relative importance of sequential and positional associations in serial learning, Jensen and Rohwer (1965) compared transfer from a serial list to two different types of paired-associate tasks. For one design, subjects learned a serial list (say, A, B, C, D, . . .) and then a double-function list of paired associates, a list constructed from adjacent items of the serial task (A-B, B-C, C-D, . . .). The control subjects learned a serial list of items unrelated to those on the list of paired associates. This experimental arrangement was viewed as a test of the *sequential hypothesis*, according to which serial learning consists of the formation of associations between successive pairs of items in a list.

For the other type of design, Jensen and Rohwer's subjects learned a serial list (A, B, C, D, . . .) and then were given a paired-associate task that required them to associate individual items of the serial list with spatial positions in a horizontal array of rectangles (1-A, 2-B, 3-C, 4-D, . . . , where each number represents a spatial position in the array). When a red dot appeared in one of the rectangles, subjects were to respond with the item that had occupied the corresponding ordinal position in the serial list. A control group learned the same spatial paired-associate task after having learned a serial list constructed of items unrelated to that transfer task. This design was conceived as a test of the *ordinal-position hypothesis*, according to which serial learning consists of the formation of associations between individual items and their respective ordinal positions in the list.

Jensen and Rohwer assumed that the differences in transfer effects

yielded by the two designs would reflect the relative availability of sequential versus positional associations after serial learning and, in turn, reflect the extent to which the formation of associations of each type is involved in serial acquisition. The results, however, failed to provide any evidence that could be taken as support for either the sequential hypothesis or the ordinal-position hypothesis. Not only were there no differences between experimental and control conditions as a function of type of paired-associate (transfer) task, but neither paradigm resulted in appreciable transfer effects. Jensen and Rohwer were forced to conclude that neither sequential nor positional associations appear to play a very important role in serial learning.

A number of subsequent studies have yielded significant transfer effects with the serial/double-function design. However, it has become clear that a major factor determining the degree of transfer with this paradigm is the length of the anticipation interval used for the paired-associate learning (Heaps, Greene, and Cheney, 1968). With few exceptions (Postman and Stark, 1967), the studies showing significant transfer effects have used a slower presentation rate for the paired associates than the 2:2-sec rate used by Jensen and Rohwer. Similarly, significant transfer effects have been reported for the serial/spatial-discrimination paradigm where a 3:3-sec rate was used (Ebenholtz, 1963), although the degree of such effects may have been inflated by the nature of the control condition employed.

It seems quite likely, therefore, that differences in relative transfer effects yielded by the two types of design may have been obscured in the Jensen and Rohwer study by the low degree of absolute transfer that was imposed by the rapid presentation of the transfer (paired-associate) lists. In view of this possibility, the essential aspects of the Jensen and Rohwer design were replicated in the present study. In order to make conditions more favorable for transfer, a 4:4-sec rate was used for presentation of those lists.

Also, the design was extended by including negative transfer conditions for both the sequential task and the positional task. In addition, a third, sequential/positional, type of transfer task was administered to groups of positive, negative, and control subjects. This task was constructed so that performance might be sensitive to the transfer of either sequential associations or positional associations as well as to a combination of the two. That is, subjects in the positive transfer condition got a positional cue and a sequential one. This combined task was included in the design to investigate the possibility that acquisition of a serial list represents a multiple-cuing type of learning for which the

'effective stimulus' is actually a *complex* of stimuli consisting of some combination of positional and sequential cues. This notion implies that maximum transfer effects would be obtained on the sequential/positional task relative to the sequential or positional tasks. The design thus included three conditions of transfer (positive, negative, and control) for each of three types of paired-associate transfer tasks (sequential, positional, and sequential/positional).

METHOD

—Materials—Subjects learned one of three *serial* lists, each consisting of a starting symbol (a set of asterisks) and ten two-syllable adjectives. The first list (S_1) contained these items: ready, basic, vocal, timid, erect, adept, overt, funny, quiet, upper. The second serial list (S_2) consisted of the same items but arranged in such a way that none occupied the same position across lists and none either followed or preceded another given item in both lists. Subjects in positive or negative transfer conditions learned S_1 or S_2 . Subjects in the control conditions learned a third serial list (S_3), which contained these items: aware, exact, woven, giant, outer, crazy, tired, pious, handy, legal.

For the *sequential* transfer tasks, two nine-item, double-function lists of paired associates (sequential-association lists SA_1 and SA_2) were derived, one from serial list S_1 and one from S_2 . Individual pairs were constructed from items occupying adjacent positions in the appropriate serial list. Subjects in the positive condition learned either S_1 and SA_1 , or S_2 and SA_2 . Subjects in the negative condition learned S_1 and SA_2 , or S_2 and SA_1 . Control subjects learned either S_3 and SA_1 , or S_3 and SA_2 .

For the *positional* transfer tasks (here termed OP_1 and OP_2), there were constructed horizontal arrays of ten rectangles, with each rectangle to be associated with one of the ten items appearing in serial lists S_1 and S_2 . From each of the lists S_1 and S_2 , a task was derived so that for positive transfer conditions (S_1 and OP_1 , or S_2 and OP_2), the arrangement of the items across the rectangles corresponded to the ordinal positions of the items in the appropriate serial list. For negative transfer conditions (S_1 and OP_2 , or S_2 and OP_1), the position of each item in the spatial array was different from that in which the item occurred in the serial list. For the control conditions (S_3 and OP_1 , or S_3 and OP_2), items in the spatial task were unrelated to those in serial list.

For the *sequential/positional* tasks, a transfer list again was derived from each of the serial lists S_1 and S_2 . This transfer task combined the sequential and positional tasks. That is, for each response item a 'compound stimulus' was presented that consisted of the position of a red dot in the array as well as another of the items in the transfer list. The stimulus word of a given pair appeared simultaneously with the position cue and in the rectangle to the immediate left of the one containing the red dot. This task consisted of one less item than did the positional task, since the first rectangle was used only to present the stimulus word for the item occupying the second

spatial position. For the positive conditions here (S_1 and $SA-OP_1$, or S_2 and $SA-OP_2$), the position of a given item in the spatial array corresponded to the position of the item in the serial list, and the word that appeared as a stimulus was the one that immediately preceded a given response item in the serial list. For the negative conditions (S_1 and $SA-OP_2$, or S_2 and $SA-OP_1$), response items of the transfer task occupied spatial positions that were different from the ordinal positions held by the items in the serial task. Further, stimulus and response terms for each of the nine pairs represented items in nonadjacent positions on the serial list. The lists for the control conditions (S_3 and $SA-OP_1$, or S_3 and $SA-OP_2$) contained items unrelated to the serial task. All lists were typed in uppercase letters and were presented by means of an MTA Scholar.

—Subjects—With two transfer lists for each condition of transfer (positive, negative, and control) under each type of task (sequential, positional, and sequential/positional), the design included 18 groups of subjects. One hundred and sixty-two college students, who volunteered to participate, served as subjects. Nine subjects were assigned to each of the 18 groups. Assignment was carried out in blocks of 18 with a different predetermined random order of the 18 conditions being used for each block. Three experimenters collected the data, each using an independent assignment schedule and running one-third of the subjects in each condition.

—Procedure—After receiving standard instructions for serial anticipation learning, the subjects were presented one of the three serial lists at a 3-sec rate with a 3-sec intertrial interval. Training on the serial list was to a criterion of two successive perfect trials at the task. Immediately after serial training, subjects in positive transfer conditions were given explicit information about the transfer task to be learned and the nature of the interlist relationship. Instructions were read in conjunction with the presentation of a card depicting the appropriate transfer paradigm in terms of geometric symbols of different colors. Subjects in the negative transfer conditions were similarly instructed, except that they were informed of the inappropriateness of first-list positional and/or sequential relations to the transfer task. Control subjects were given information about the construction of the relevant transfer list and told that it was unrelated to the serial list. The duration of the intertask interval was approximately 5 min.

For the sequential transfer task, each successive stimulus word appeared for 4 sec, followed by the joint appearance of stimulus and response terms for an additional 4 sec. For the positional transfer task, a red dot representing a given positional cue was presented in one of a series of ten rectangles arranged horizontally across the window of the drum. After a 4-sec anticipation interval, the drum turned to present the appropriate response word positioned in the array in the place of the red dot. Then 4 more sec elapsed before the appearance of the positional cue for the next item. For the sequential/positional transfer list, a red dot was presented in one of the rectangles in conjunction with the appropriate stimulus word, which appeared in the 'box' immediately to the left of the one containing the red dot. After 4 sec the drum turned to reveal both the stimulus word and the response word placed in the appropriate rectangles of the array. After

4 sec the pair of cues for the next item appeared. A 4-sec intertrial interval was used for all these paired-associate lists, and each list was presented in three different random orders so as to reduce serial effects on the paired-associate learning here. Subjects were given ten trials on the transfer task, with no responses being required on the first presentation.

RESULTS

Serial performance

Performance on the serial task was compared across the nine groups representing the factorial arrangement of three levels of the task variable with three conditions of transfer. Mean numbers of trials to achieve two successive errorless recitations of the list ranged from 9.72 to 12.00, the differences among groups being nonsignificant [$F < 1$, $df = 8, 153$].

Early transfer effects

Performance on the first anticipation trial of paired-associate learning was used to assess early transfer effects. Shown in Table 1 are mean numbers of first-trial errors for positive, negative, and control conditions for each of the three tasks. Analysis of variance indicated significant effects both for tasks ($F = 5.37$, $df = 2, 153$; $p < .01$) and for conditions of transfer ($F = 87.28$, $df = 2, 153$; $p < .001$). The interactive effects of the two factors were also significant ($F = 2.67$, df

Table 1. Mean numbers of errors on the first transfer trial for each condition on each task

Task	Condition		
	Positive	Control	Negative
Positional	1.94	6.44	6.22
Sequential/positional	1.28	6.33	6.44
Sequential	3.94	6.50	6.89

= 4, 153; $p < .05$). Pairwise comparisons among the three tasks with the Newman-Keuls test revealed that while the positional and sequential/positional tasks did not differ from each other, each was superior to the sequential task ($p < .05$). Positive transfer effects were highly reliable ($p < .005$ by Dunnett's test), while first-trial performance yielded no indication of negative transfer. The significant interaction

between tasks and conditions of transfer reflects a greater amount of positive transfer for the positional and sequential/positional tasks than for the sequential task. Differences between positive and control conditions for the positional and sequential/positional tasks (combined) contrasted with those for the sequential task were significant at the .005 level ($F = 9.56, df = 1, 104$). Differences in the amount of positive transfer for positional versus sequential/positional tasks were not significant ($F < 1, df = 1, 68$). According to Murdock's (1957) formula, $(C - P)/(C + P) \times 100$, applied to total numbers of first-trial errors for positive and control conditions, percentages of positive transfer were 53.64, 66.42, and 24.47 for the positional, sequential/positional, and sequential tasks respectively.

Overall transfer effects

Performance curves for the nine experimental groups are presented in Figure 1 in terms of mean number of errors over the nine transfer trials in blocks of three trials each. These curves suggest considerable

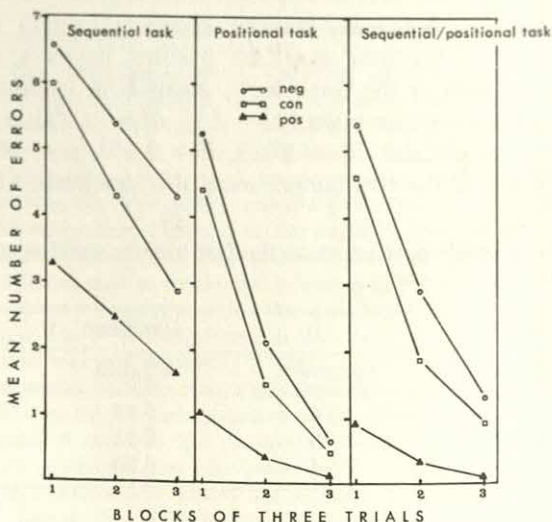


Fig. 1. Mean number of errors for positive, negative, and control conditions on each of three transfer tasks

transfer effects within each task, in addition to reflecting differences in overall level of difficulty across the three types of task. Analysis of

variance applied to number of errors on the transfer lists indicated significant effects both for task ($F = 55.22$, $df = 2, 153$; $p < .001$) and for condition of transfer ($F = 42.51$, $df = 2, 153$; $p < .001$). Interactive effects of task and transfer condition were nonsignificant ($F < 1$, $df = 4, 153$). Mean numbers of errors for the sequential, positional, and sequential/positional tasks were 36.76, 15.30, and 18.28 respectively. Pairwise comparisons with the Newman-Keuls test indicated significant differences ($p < .01$) between the sequential and positional tasks and between the sequential and sequential/positional ones. The difference between the positional and sequential/positional tasks did not approach significance. For conditions of transfer, mean numbers of errors were 10.07, 26.54, and 33.72 for positive, control, and negative respectively. Dunnett's test showed that both positive and negative transfer effects were significant beyond the .005 level.

Since level of difficulty varied markedly as a function of type of task, absolute differences in number of errors among positive, negative, and control conditions for the three tasks are not likely to provide a very sensitive measure of the *relative* degree of transfer for each type of task (Jensen and Rohwer, 1965). That is, a given difference between positive versus control conditions for the positional or sequential/positional tasks may represent a greater proportion of transfer than for the sequential one because of a ceiling effect imposed by the relative lack of difficulty of the other two tasks.

So in order to provide an explicit test of relative transfer effects for the three types of tasks, the following analyses were conducted. The 18 subjects within each of the nine experimental conditions were ranked on the basis of number of errors to first-list criterion. For each type of task, transfer data were arranged in 18 blocks such that each block represented error scores of three subjects (one from each transfer condition) matched according to their ranks on first-list performance. For each of the 18 blocks (within the three tasks), percentages of positive and negative transfer were calculated on the basis of scores of the three matched subjects within a block. Percent positive transfer was computed according to the formula $(C - P)/(C + P) \times 100$, where C and P represent the scores of a control subject and a positive-transfer subject respectively. Similarly, percent negative transfer was calculated on the basis of $(C - N)/(C + N) \times 100$, where N represents the score of a negative-transfer subject. In order to determine whether differences in relative transfer across the three tasks were related to the positions the items occupied in the serial list, percent transfer was calculated separately for items taken from the beginning, middle, and end

portions of the serial list. For the positional task, these portions corresponded to the items in positions 1-4, 5-7, and 8-10, respectively, of the serial list. For the sequential and sequential/positional tasks, the beginning items were 2-4 of the serial task, since the item appearing in position 1 was used only as a stimulus term in the transfer task.

Presented in Table 2 are mean percentages of positive transfer for the three tasks as a function of serial positions of response items. Each of the entries represents the average of 18 values calculated individually for matched pairs of subjects. These data reveal substantial differences in relative transfer effects across the three types of transfer task. Mean percentages of positive transfer (averaged over serial positions) were 62.33 for the positional, 73.43 for the sequential/positional, and 36.00 for the sequential task. Analysis of variance of these percentages (transformed to arcsin) indicated a significant effect for type of transfer task ($F = 9.84$, $df = 2, 51$; $p < .001$). Pairwise comparisons among the three means by the Newman-Keuls test indicated significantly greater transfer for the positional and sequential/positional tasks than for the sequential task ($p < .01$). The difference between the positional and sequential/positional tasks was not significant. Percent positive transfer also varied significantly as a function of the position of items in the serial list ($F = 4.31$, $df = 2, 102$; $p < .025$); the mean percent-

Table 2. Mean percentages of positive transfer by task and serial position

Task	Serial position		
	Beginning	Middle	End
Positional	77.87	49.78	59.33
Sequential/positional	74.82	77.41	68.07
Sequential	40.18	35.91	31.92

Table 3. Mean percentages of negative transfer by task and serial position

Task	Serial position		
	Beginning	Middle	End
Positional	14.58	2.58	16.02
Sequential/positional	8.58	9.62	13.77
Sequential	11.66	8.90	15.03

ages for beginning, middle, and end positions (averaged across tasks) were 64.29, 54.37, and 53.11 respectively.

A significant portion of the variance among percent positive transfer

scores was attributable to the interactive effects of task and serial position ($F = 2.53$, $df = 4, 102$; $p < .05$). Analysis of simple effects showed reliable differences among the three tasks at each level of the serial-position variable, F in each case ($df = 2, 51$) being significant beyond .005. On items taken from either the first or the last portion of the serial list, relative positive transfer for the sequential task was less than that for either of the other two tasks ($p < .01$ by the Newman-Keuls test), with no significant differences between the latter two. However, on those items taken from the middle positions of the serial list, relative transfer effects for the positional task were significantly less pronounced than those for the sequential/positional task ($p < .05$) and did not differ significantly from those yielded by the sequential task.

Mean percentages of negative transfer for each of the three tasks as a function of serial position of response terms are shown in Table 3. Negative transfer effects for each task were markedly lower in magnitude than positive transfer effects. Analysis of variance applied to arcsin transforms of percent negative transfer scores yielded no significant effects. Mean percentages for positional, sequential/positional, and sequential tasks respectively (averaged across serial positions) were 11.06, 10.66, and 11.86. The degree of negative transfer yielded by the positional task on items taken from the middle of the serial list was notably less than that on items taken from either the first or the last portion. However, the magnitude of this effect was not sufficient to produce a significant interaction of tasks and positions ($F < 1$, $df = 4, 102$).

DISCUSSION

The present results indicate that when subjects are given explicit instructions as to the interlist relationship and when the transfer list is presented at a relatively slow rate, both positive and negative transfer effects may be obtained in the serial/paired-associate paradigm. Such effects were apparent for the type of design which has been assumed to reflect the transfer of sequential associations as well as for that which has been proposed as a measure of the transfer of positional associations from serial to paired-associate learning.

A comparison of the present results with those of Jensen and Rohwer (1965) emphasizes the importance of allowing the subject sufficient time to utilize the cues provided in a paired-associate list to mediate

transfer from the serial list. Heslip (1968) has cautioned against the use of an anticipation interval in excess of 3 sec, on the grounds that the subject may, rather than responding directly to the cue presented in paired-associate learning, covertly recite the sequence of items before emitting the appropriate response. Since the 4-sec anticipation interval used for presentation of the paired associates in the present study was a little over the 3-sec limit suggested by Heslip, it is possible that absolute transfer effects were inflated to some degree by this factor. However, it should be noted that *differential* transfer effects across the three tasks are independent of any such 'recitation confounding,' since the same anticipation interval was used in the presentation of the three tasks' lists of paired associates.

If the difference in degree of relative positive transfer yielded by the sequential and positional tasks is taken as an indication of the relative importance of sequential versus positional associations in serial acquisition, the results appear to favor the ordinal-position hypothesis. The amount of positive transfer on the first trial of paired-associate learning for the positional task was clearly superior to that for the sequential task. Overall positive transfer effects, when adjusted for differences in task difficulty, were almost twice as pronounced in the positional than in the sequential task.

Although negative transfer effects over the three tasks (combined) were significant, they were of a much lower order than positive transfer effects, and they failed to distinguish among the three types of serial/paired-associate paradigms. The high degree of interlist distinctiveness between the serial task and the transfer tasks quite likely served to reduce the strength of potential interlist-intrusion tendencies.

Another factor that may have contributed to the asymmetry of positive and negative transfer effects (at least early in training) is the differential response availability at the beginning of paired-associate learning for subjects in positive or negative transfer versus control conditions. Since control subjects were required to learn a set of responses different from those used in the serial task, comparison of their paired-associate performance to that of positive-transfer subjects is likely to lead to an overestimate of positive transfer, and comparison of their performance to that of negative-transfer subjects is apt to yield an underestimate of negative transfer. Since the present study was concerned primarily with relative transfer effects, this issue is crucial only to the extent that differences between experimental and control subjects in terms of response availability varied across the three paired-associate tasks. It may be noted that control subjects in the sequential

and sequential/positional groups were given a greater opportunity to learn response items as such during paired-associate presentation than were the positional control subjects, since for the first-mentioned two groups all but one of the response terms also appeared as stimulus terms. However, the almost identical curves for control subjects in the positional and sequential/positional groups (see Figure 1) suggest that relative transfer effects were not confounded by differential rates of response acquisition for control subjects under the various task conditions.

Differences in the amount of relative transfer for the sequential and positional tasks as a function of the position of paired-associate items in the serial list may be related to the 'dual process' hypothesis (see Jensen and Rohwer, 1965). According to the most recent version of this hypothesis, positional cues play a major role in the acquisition of items at the ends of a serial list, while sequential cues serve as the primary 'functional stimuli' for items in the middle of the list. In the context of the present experiment, this notion suggests that for paired-associate items taken from the ends of the serial list, transfer with the positional task should be greater than that with the sequential one; while for items taken from the middle of the serial list, transfer with the sequential task should be superior to that with the positional task. But the results indicated no tendency for relative transfer with the sequential task to be greater than that with the positional task—regardless of serial position of response items. That transfer effects for the positional task on items taken from the middle portion of the serial list were less than those on items taken from the ends does suggest a decreased *effectiveness* of positional cues with respect to middle items vis-à-vis end items. However, the conclusion that the stimulus function in the acquisition of middle items is provided primarily by sequential cues seems to be precluded, since relative positive transfer was still appreciably greater on middle items in the positional than in the sequential condition.

A number of factors lead to an expectation of superior transfer effects for the sequential/positional task relative to the sequential or positional ones. First, if positional cues function as stimuli at one portion of a serial list and sequential cues at another, transfer effects should be maximum where both types of cues are presented in the transfer task. Second, if there are individual differences across subjects in their preference for 'strategies' involving positional versus sequential learning, the sequential/positional task should be favored. Third, if sequential and positional cues 'overlap' to form a complex of stimuli

appropriate to each item in the serial list, the degree of generalization decrement between serial and paired-associate tasks should be less for the sequential/positional task than for the sequential or positional tasks.

Relative positive transfer for the sequential/positional task was grossly superior to that for only the sequential task, however; and the degree of this superiority was fairly constant regardless of the positions in which transfer-list items appeared in the serial list. It seems likely that subjects in the sequential/positional condition refrained from using sequential cues to any appreciable extent to mediate transfer performance. The similarity of the results for the sequential/positional and positional tasks, both in terms of relative positive transfer and overall level of performance, further suggests that the sequential cues provided in the sequential/positional task were not a major factor determining the subjects' responses. The tendency of these subjects to select the positional rather than the sequential cues to mediate transfer may indicate that during serial learning the strength accrued by positional associations is greater than that accrued by associations between adjacent items.

It may be argued, of course, that the subjects' predominant use of positional cues reflects a strategy of 'ignoring' the sequential aspects of the sequential/positional task in order to avoid the rather extreme difficulty inherent in a double-function list (see Postman and Stark, 1967). Since all but two of the items in the sequential or sequential/positional tasks served both a stimulus and a response role, heavy interference from backward associations would provide a major source of difficulty in paired-associate acquisition. Subjects in the sequential/positional task might eliminate this difficulty by responding strictly on the basis of the positional cues presented. However, the greater amount of relative positive transfer for the sequential/positional task than for the positional task on items taken from the middle portion of the serial list shows that subjects did not completely disregard the sequential cues in the sequential/positional task.

For items taken from the ends of the serial list, asymptotic transfer performance was achieved by the presentation of positional cues alone. However, for items taken from the middle of the serial list, the effectiveness of positional cues in mediating transfer was enhanced when they appeared in combination with relevant sequential cues. Thus, while comparison of overall relative positive transfer across the three paired-associate tasks suggests that interitem associations constitute a less essential factor in serial learning than do positional associations, the

results are not incompatible with the assumption that sequential cues form a part of the complex serving as 'functional stimulus' for at least some of the items in a serial list.

Notes

This research was supported by a grant from the National Research Council of Canada, APA 0276. Received for publication July 1, 1971; revision, August 11, 1971.

References

- Ebenholtz, S. M. 1963. Position mediated transfer between serial learning and a spatial discrimination task. *Journal of Experimental Psychology* 65:603-608.
- Heaps, R. S., Greene, W. A., and Cheney, C. D. 1968. Transfer from serial to paired-associate learning with two paired-associate rates. *Journal of Verbal Learning and Verbal Behavior* 7:840-841.
- Heslip, J. R. 1968. A methodological note on direct tests of the effectiveness of positional cues in serial learning. *Psychonomic Science* 13:221-222.
- Jensen, A. R., and Rohwer, W. D., Jr. 1965. What is learned in serial learning? *Journal of Verbal Learning and Verbal Behavior* 4:62-72.
- Murdock, B. B., Jr. 1957. Transfer designs and formulas. *Psychological Bulletin* 54:313-326.
- Postman, L., and Stark, K. 1967. Studies of learning to learn: IV. Transfer from serial to paired-associate learning. *Journal of Verbal Learning and Verbal Behavior* 6:339-353.

continued from page 462

- Kimmel, H. D., ed. *Experimental Psychopathology: Recent Research and Theory*. New York: Academic Press, 1971. Pp. xiii, 264. \$12.50.
- Kittrie, Nicholas, N. *The Right to Be Different*. Baltimore: The Johns Hopkins Press, 1971. Pp. xxii, 443. Paperback, \$4.95.
- Klein, Stuart M. *Workers under Stress: The Impact of Work Pressure on Group Cohesion*. Lexington: University Press of Kentucky, 1971. Pp. xii, 123. \$7.50.
- Klingender, Francis. *Animals in Art and Thought to the End of the Middle Ages*. Cambridge, Mass.: M.I.T. Press, 1971. Pp. xxviii, 580.
- Klinger, E. *Structure and Functions of Fantasy*. New York: Wiley, 1971. \$13.50.
- Knox, David. *Marriage Happiness: A Behavioral Approach to Counseling*. Champaign, Ill.: Research Press, 1971. Pp. xi, 171. Paperback.
- Koestenbaum, Peter. *The Vitality of Death*. Westport, Conn.: Greenwood Press, 1971. Pp. xi, 589. \$15.00.
- Koppitz, Elizabeth Munsterberg. *Children with Learning Disabilities: A Five Year Follow-up Study*. New York: Grune & Stratton, 1971. Pp. xv, 218.
- Kussman, Thomas. *Bewusstsein und Handlung: Probleme und Ergebnisse der sowjetischen Psychologie*. Stuttgart: Hans Huber, 1971. Pp. 233.
- Lester, Gene, and Lester, David. *Suicide: The Gamble with Death*. Englewood Cliffs, N. J.: Prentice-Hall, 1971. Pp. xiii, 176. \$5.95; paperback, \$2.45.
- Lohr, Thomas F. *The Mechanics of the Mind*. Coopersburg, Pa.: Venture Books, 1971. Pp. xi, 516. \$11.95.
- Lorenz, Konrad. *Studies in Animal and Human Behavior*. Vol. 2. Cambridge, Mass.: Harvard University Press, 1971. Pp. xxiv, 366. \$11.75.
- McFarland, D. E., ed. *Personnel Management*. Baltimore: Penguin Books, 1971. Pp. 402. Paperback, \$3.95.
- Mackintosh, N. J., and Honig, W. K., eds. *Fundamental Issues in Associative Learning: Proceedings of a Symposium Held at Dalhousie University, Halifax, June 1968*. Halifax, Nova Scotia: Dalhousie University, 1971. Pp. 203. \$8.00.
- Malherbe-Weil, Hans, and Szara, Stephen I. *The Biochemistry of Functional and Experimental Psychoses*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xvii, 406. \$18.50.
- Maltz, Maxwell. *Power Psycho-Cybernetics for Youth*. New York: Grosset & Dunlap, 1971. Pp. 217. \$5.95.
- Martindale, Don, and Martindale, Edith. *The Social Dimensions of Mental Illness, Alcoholism, and Drug Dependence*. Westport, Conn.: Greenwood Press, 1972. Pp. xvi, 330. \$12.50.
- Maruyama, Magoroh, and Dator, James A., eds. *Human Futuristics*. Honolulu: University of Hawaii Press, 1971. Pp. xviii, 237. \$7.00.
- Mausner, Bernard, and Platt, Ellen S. *Smoking: A Behavioral Analysis*. Elmsford, N. Y.: Pergamon Press, 1971. Pp. xiii, 238. \$14.00.

continued on page 350

PERCEPTUAL LEARNING IN YOUNG CHILDREN: VARIED CONTEXT AND STIMULUS LABELS

Cynthia Turnure

University of Minnesota

The 48 nursery schoolers saw five trials of one letterlike form with a different color, size, and background on every trial (varied form) and five trials of another form with the same color, size, and background on every trial (constant form). Half got nonsense names for the two forms. On a recognition test with the two forms plus six others, fewer errors were made by older than younger and by middle- than lower-class children. Their initial choices revealed no effects of contextual variation versus constancy, or of stimulus labels, but younger subjects chose the varied and older subjects the constant form as the one they remembered best.

It has frequently been suggested (Jeffrey, 1969; Scott, 1966; Trabasso, 1968) that one way to facilitate learning and problem solving in the young or mentally retarded child is to eliminate or hold constant all stimulus cues that are not relevant to the solution of the immediate task. The empirical basis for this recommendation has primarily been discrimination-learning and concept-identification experiments, which have repeatedly demonstrated that the fewer and less variable the irrelevant cues, the faster the learning (Eimas, 1966; House and Zeaman, 1963; Lubker, 1967; Osler and Kofsky, 1965; Trabasso, 1963; Trabasso and Bower, 1968).

Several other recent studies, however (Odom and Guzman, 1970; Gibson and Shepela, unpublished manuscript; White, 1965), have shown that the young child learns just as well when irrelevant cues are varied as when they are held constant. Indeed, it has even been suggested that the very young child may learn *better* under conditions of stimulus variation (Jeffrey, 1969; Odom and Guzman, 1970), although there is disagreement on this point (Gollin, 1961; White, 1966). The present study was an attempt to specify the learning situations and children that might benefit from stimuli presented in varying versus constant contexts.

Tighe and Tighe (1966) have pointed out the importance of distinguishing between two processes in discrimination learning, each of

which may be facilitated by different conditions. The first process is perceptual learning, defined as increased sensitivity to variables present in stimulation; the second is associative learning, or the association of responses to previously discriminated variables of stimulation. The discrimination-learning and concept-identification experiments mentioned above have been concerned primarily with associative learning. In those experiments it has been assumed that the subject is able to discriminate all the task cues (say, color, size, and form) prior to his learning which of those cues was arbitrarily designated as correct for a particular problem. In addition, reinforcement or some kind of feedback on correctness of response has usually been given on each trial.

There seem to be several limitations of such studies as a basis for recommendations for enhancing the learning of young children in other than controlled experimental settings. First, often it cannot be assumed that young children can (or will) discriminate all the cues relevant to a task before the task begins. One of the primary goals of reading-readiness training, for example, is to *teach* the child to recognize the letters of the alphabet and thus be able to discriminate which aspects of letterlike forms are relevant to correctly identifying letters and which are irrelevant. Before recommending the elimination of contextual variation in classroom and other general learning situations, the effects of this variable should be examined in tasks in which the child has not previously learned all the cues—in perceptual-learning tasks, for example.

Another limitation of those associative-learning studies concerns the use of reinforcement on every trial. Again, this does not seem to parallel the typical learning situation of the young child, who continually discovers regularities in his environment *without* reinforcement or feedback on every 'trial' (see Gibson, 1969; Tighe and Tighe, 1966). One of the most obvious and significant examples of this type of early learning is the child's acquisition of language, in which he is clearly not reinforced in any traditional sense for every trial utterance or each instance of listening. Yet he learns in a most impressive way to recognize and use the correct patterns of spoken language.

Indeed, one of the most striking characteristics of language is the great variation of the context in which it is spoken (its pitch, loudness, speed, setting, and so forth; see Gibson, 1969). Thus again, before eliminating contextual variation from other early learning situations, it seems important to investigate its effects further. Contextual variation has previously been found to facilitate the differentiation of parts

from wholes by college students (Turnure and Wallach, 1965), and the present study was an attempt to extend the investigation of its effects to young children under two conditions of perceptual learning.

One other suggestion often put forth for facilitating early learning is to provide the subject with a name or label for the stimulus to be discriminated. A number of studies (Cantor, 1955; Norcross and Spiker, 1957; Stephens, 1968; Stern, 1965) have reported superior performance on discrimination and concept-learning tasks after the subject has learned names for the stimuli, and several studies have also found that labels facilitate short-term memory and delayed reaction-time responses in children (Bernbach, 1967; Bush and Cohen, 1970; Spiker, 1956). Other studies, however, have failed to find any difference between the effects of learning labels and other types of stimulus experience (such as observation or stimulus judgments) on subsequent discrimination performance by adults (Arnoult, 1953; Goss, 1953; Pfafflin, 1960; Rasmussen and Archer, 1961; Robinson, 1955; Turnure and Wallach, 1965). In addition to contextual variation, then, the present study also investigated the effects of presence versus absence of stimulus labels on perceptual learning by young children.

METHOD

—Subjects—The subjects were 48 children attending the University of Minnesota Nursery School. There were 24 in an *older* group (12 boys and 12 girls), ranging in age from four years and five months to five years and seven months (mean age, five years), and 24 (14 boys and 10 girls) in a *younger* group, with ages ranging from three years to four years and four months (mean age, three years and ten months). Most of the children came from homes that could be characterized as upper middle class in socioeconomic status, with fathers employed primarily in the professions, but there were also five children in each age-group who come from homes of lower socioeconomic status, homes located in a predominantly black Model Cities area. These children's families' incomes were below the Headstart guidelines, and the children attended the school tuition-free.

—Experimental design and stimuli—The children viewed a series of slide presentations of letterlike forms, seeing five 5-sec presentations of one form in which the color of the form, the color of the background, and the size of the form were different on every trial (*varied form*). Each child also saw five trials of a second form that had the same color, size, and background on every trial (*constant form*).

Table 1 shows which training stimuli were seen by the 12 subjects in each of the four experimental groups. The training stimuli were modeled after two of Pick's (1965) letterlike forms: $\wedge$ and $\oslash$, referred to here as A and B. Six versions of the two forms, each version with a different color, size, and

Table 1. Training stimuli for subjects in each experimental group

Group (age/condition)	Sub- ject	Constant form			Varied forms			
		Color	Back- ground	Size (in.)				
Older/No-name (N = 12) <i>or</i> Older/Name (N = 12) <i>or</i>	1	A ₁	green	14	B ₂	B ₃	B ₄	B ₆
	2	A ₂	red	11	B ₁	B ₃	B ₄	B ₆
	3	A ₃	blue	6	B ₁	B ₂	B ₄	B ₆
	4	A ₄	yellow	4½	B ₁	B ₂	B ₃	B ₆
	5	A ₅	purple	3	B ₁	B ₂	B ₃	B ₆
	6	A ₆	orange	2½	B ₁	B ₂	B ₃	B ₅
Younger/No-name (N = 12) <i>or</i> Younger/Name (N = 12)	7	B ₁	green	14	A ₂	A ₃	A ₄	A ₆
	8	B ₂	red	11	A ₁	A ₃	A ₄	A ₆
	9	B ₃	blue	6	A ₁	A ₂	A ₄	A ₆
	10	B ₄	yellow	4½	A ₁	A ₂	A ₃	A ₆
	11	B ₅	purple	3	A ₁	A ₂	A ₃	A ₆
	12	B ₆	orange	2½	A ₁	A ₂	A ₃	A ₅

background, were prepared as slides. Each of the 12 subjects in each of the four experimental groups viewed a different one of these 12 training stimuli as his constant form. Each child also saw five different versions of the other letterlike form as his varied form (see Table 1). The order of stimulus presentation for each child was determined randomly, with the restriction that no child see more than three A or B forms in a row.

Half of the subjects in each age-group were given one nonsense name on each trial of their varied form and a second name on each trial of their constant form. Two CVC trigrams (wug and mib), matched for meaningfulness (Archer, 1960), served as names, with half the subjects in the *name condition* receiving wug as the name of their varied form and mib for their constant form, and half receiving mib for their varied form and wug for their constant form. None of the subjects in the *no-name condition* were told any name for either form.

The stimuli for the recognition test were eight of Pick's (1965) letterlike forms, appearing together on one slide as black forms on a white background. This slide included the two training forms plus six forms the subjects had not previously seen; here all forms were approximately 2 in. in size.

—Procedure—Each child was tested individually in a small experimental room. The child was seated on a small chair approximately 3 ft in front of a projection screen. The training and test slides were projected on the screen by a Kodak Carousel slide projector. Each training slide was shown for 5 sec. The test slide remained on the screen as long as necessary to obtain the child's responses.

The subjects in the no-name condition were given these training instructions: "I'm going to show you some pictures. I want you to look very carefully at each one. Afterwards I'm going to ask you about what you saw. So look carefully at each slide." The subjects in the name condition were given the same instructions but were also told: "I will also tell you the name of each picture." During training in the name condition the experimenter then told the child the name of each stimulus while it was being shown ("This is a wug," for example).

After the ten training slides had been presented, the recognition test slide was shown. Each child was asked, "Which of these forms do you think you just saw in the other slides?" If he didn't choose immediately, he was asked, "Which of these do you remember from the slides you just saw?" If he chose only one form he was asked, "Do you remember any of the other ones?" If he chose more than one form, he was asked, "Which one do you remember *best*?" Finally, at the end of the experiment, each child got to choose a small toy as a "prize for playing the game."

RESULTS

The recognition-test choices were analyzed using nonparametric (χ^2 and Fisher exact-probability) tests. In addition, the error scores (choices of test forms other than the two training forms) were analyzed

by *t* tests. All significance levels reported are based on two-tailed tests.

Table 2 presents the initial responses of the groups on the recognition test. The main effects evident in these data involve differences between the age-groups. The older children were more likely to say they had seen only their two training forms before, whereas the younger ones were more likely to say they had seen those two forms plus some or all of the other test figures ($\chi^2 = 4.84, p < .05$, with 1 *df*). This age effect was also apparent in the number of errors made by the two groups: the younger subjects made significantly more errors on the recognition test than did the older ones ($t_{46} = 3.03, p < .01$).

Table 2. Initial choices on recognition test

Group (age/condition)	Varied form only	Constant form only	Both training forms only	Varied plus others
Older/No-name	0	0	10	0
Older/Name	0	0	10	0
Younger/No-name	1	0	4	0
Younger/Name	1	0	7	0
	Constant plus others	Both plus others	Neither training form	Errors (all subjects)
Older/No-name	1	1	0	4
Older/Name	2	0	0	4
Younger/No-name	1	5	1	37
Younger/Name	0	3	1	15

The error data also revealed that the ten subjects from lower socioeconomic backgrounds made significantly more errors than the predominantly upper-middle-class subjects ($t_{46} = 4.82, p < .01$). In order to determine whether or not the errors of these ten children were primarily responsible for the significant age effect in the error data, the scores of only the upper-middle-class subjects were analyzed. Within this more homogeneous sample, the younger subjects still made significantly more errors than the older subjects on the recognition test ($t_{36} = 2.28, p < .05$). Further analysis of the error scores of only the ten subjects with lower socioeconomic status indicated that the age effect was significant for that subsample as well ($t_8 = 2.80, p < .05$).

It should be noted that the initial responses on the recognition test

revealed no significant differences in the subjects' recognition of their varied versus their constant training form. Nor were there any significant differences between subjects in the name versus the no-name conditions, either in the number of children who recognized one or both training forms or in the number of errors on the recognition test.

The choices of the subjects when asked which test form they remembered best are shown in Table 3. Analysis of these data revealed a significant age effect similar to the one evident in the initial-response data. On this part of the recognition test, the older subjects were more likely to say they remembered both of their training forms best, whereas the younger subjects tended to say they remembered neither training form best, or another of the test forms best ($p < .05$ by the Fisher exact-probability test). These data also revealed a significant interaction between age and type of stimulus training. Of those children who said they remembered one training form best, the younger ones tended

Table 3. Form remembered best on recognition test

Group (age/condition)	Varied form	Constant form	Both training forms	Neither training form	Another test form
Older/No-name	5	4	2	0	1
Older/Name	0	5	6	1	0
Younger/No-name	5	0	1	2	4
Younger/Name	3	1	3	2	3

to remember their varied form best whereas the older ones tended to remember their constant form best ($p < .05$ by the Fisher exact test). Finally, as in the initial-response data, there were no significant effects associated with the name versus no-name conditions in the children's choices of the form they remembered best.

DISCUSSION

The initial recognition-test responses in the present study revealed no significant effects of contextual variation versus constancy. This finding suggests that in perceptual-learning situations like that of the present study, there may be no particular advantage to "impoverishing the environment" (Trabasso, 1968) by minimizing irrelevant cues, at least as far as the children's immediate memory for stimuli is con-

cerned. As Fellows (1968) has pointed out, there are many other ways of directing attention to the stimuli that must be learned—without impoverishing the learning environment.

The findings on which form was best remembered provide some support for the view that younger children (in this study, from three to four years and three months) may be particularly responsive to situations involving stimulus variation. Odom and Guzman (1970) also found that their younger subjects (five-year-olds) learned better under conditions of stimulus variability than of constancy, and Jeffrey (1969) and Gibson (1969) have also suggested that high contextual variation may help the young child to isolate the invariant properties of objects. It is interesting that the highly successful educational television show "Sesame Street" makes extensive use of the method of varying stimulus cues (such as size, color, and background) in presenting letter and number concepts to preschoolers.

The fact that the younger children also made a number of errors on the recognition test suggests that the task did require some perceptual learning on their part, while the relatively few errors made by the older children suggests that they were able to discriminate the training and test forms fairly easily. It may be that when perceptual learning must occur (when the child cannot already discriminate the distinctive features of the stimuli), presenting the stimuli in varying contexts is helpful, whereas when the child can readily discriminate the stimuli, repeatedly presenting them in a constant context is just as effective, and perhaps even more effective.

The finding that the children from lower socioeconomic backgrounds made more errors on the recognition test than the rest of the children in this predominantly upper-middle-class sample parallels other findings on social class and perceptual-discrimination performance (Clark and Richards, 1966; Covington, 1967). This finding may be accounted for, at least in part, by the more extensive experience of the upper-middle-class children with stimuli and situations similar to those of the present experiment. It should be noted that the children of lower socioeconomic status appeared to be less attentive to the experimental task and stimuli, and much more interested in the slide projector and other nontask aspects of the situation.

Finally, providing the children in the present study with verbal labels for stimuli did not significantly affect their later recognition of those stimuli. As was pointed out in the introduction, a number of previous studies have found that stimulus labels facilitated children's discrimination learning and memory for stimuli. Several differences between these

studies and the present one may help to explain the discrepancies between the findings, however. As was noted, the present study did not provide any kind of reinforcement or feedback during training. Also, the subjects were not required to verbalize the stimulus labels on each trial, although some of them did so spontaneously. Finally, the labels used were not particularly relevant to the stimuli to be learned, because nonsense syllables were used in an attempt to parallel the situation of the child's learning names for letters.

Notes

The author thanks the staff of the University of Minnesota Nursery School for their cooperation, Anne D. Pick for permission to use stimulus materials and for a critical reading of the manuscript, Mervyn Bergman for help in preparing slides, and Marianne Kerr for help in running subjects. The author is at the Institute of Child Development, University of Minnesota, Minneapolis, Minnesota 55455. Received for publication June 28, 1971; revision, August 13, 1971.

References

- Archer, E. J. 1960. Re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychological Monographs* 74(10, whole no. 497).
- Arnoult, M. D. 1953. Transfer of predifferentiation training in simple and multiple shape discrimination. *Journal of Experimental Psychology* 45:401-409.
- Bernbach, H. A. 1967. The effect of labels on short-term memory for colors with nursery school children. *Psychonomic Science* 7:149-150.
- Bush, E. S., and Cohen, L. B. 1970. The effects of relevant and irrelevant labels on short-term memory in nursery school children. *Psychonomic Science* 18:228-229.
- Cantor, G. N. 1955. Effects of three types of pretraining on discrimination learning in preschool children. *Journal of Experimental Psychology* 49:339-342.
- Clark, A. D., and Richards, C. J. 1966. Auditory discrimination among socially disadvantaged and non-disadvantaged preschool children. *Exceptional Children* 33:259-262.
- Covington, M. V. 1967. Stimulus deprivation as a function of social class membership. *Child Development* 38(2):607.
- Eimas, P. D. 1966. Effects of overtraining, irrelevant stimuli, and training task on reversal discrimination learning in children. *Journal of Experimental Child Psychology* 3:315-323.
- Fellows, B. J. 1968. *The discrimination process and development*. Oxford: Pergamon Press.
- Gibson, E. J. 1969. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts.

- Gibson, E. J., and Shepela, S. Some effects of redundant stimulus information on learning to identify letters. Unpublished manuscript, Cornell University.
- Gollin, E. S. 1961. Tactual form discrimination: Developmental differences in the effects of training under conditions of spatial interference. *Journal of Psychology* 51:131-140.
- Goss, A. E. 1953. Transfer as a function of type and amount of preliminary experience with task stimuli. *Journal of Experimental Psychology* 46:419-428.
- House, B. J., and Zeaman, D. 1963. Miniature experiments in the discrimination learning of retardates. In *Advances in child development and behavior*, vol. 1, ed. L. P. Lipsett and C. C. Spiker, pp. 313-374. New York: Academic Press.
- Jeffrey, W. E. 1969. Early stimulation and cognitive development. In *Minnesota symposia on child psychology*, vol. 3, ed. J. P. Hill, pp. 46-67. Minneapolis: University of Minnesota Press.
- Lubker, B. J. 1967. Irrelevant stimulus dimensions and children's performance of simultaneous discrimination problems. *Child Development* 38:119-125.
- Norcross, K. J., and Spiker, C. C. 1957. The effects of type of stimulus pretraining on discrimination performance in preschool children. *Child Development* 28:179-184.
- Odom, R. D., and Guzman, R. D. 1970. Problem solving and the perceptual salience of variability and constancy: A developmental study. *Journal of Experimental Child Psychology* 9:156-165.
- Osler, S. F., and Kofsky, E. 1965. Stimulus uncertainty as a variable in the development of conceptual ability. *Journal of Experimental Child Psychology* 2:264-279.
- Pfafflin, S. M. 1960. Stimulus meaning in stimulus predifferentiation. *Journal of Experimental Psychology* 59:269-274.
- Pick, A. D. 1965. Improvement of visual and tactual form discrimination. *Journal of Experimental Psychology* 69:331-339.
- Rasmussen, E. A., and Archer, E. J. 1961. Concept identification as a function of language pretraining and task complexity. *Journal of Experimental Psychology* 61:437-441.
- Robinson, J. S. 1955. The effect of learning verbal labels for stimuli on their later discrimination. *Journal of Experimental Psychology* 49:112-115.
- Scott, K. G. 1966. Engineering attention: Some rules for the classroom. *Education and Training of the Mentally Retarded* 1:125-129.
- Spiker, C. C. 1956. Stimulus pretraining and subsequent performance in delayed reaction experiments. *Journal of Experimental Psychology* 52:107-111.
- Stephens, W. E. 1968. Concept learning with three types of verbal labels. *Psychonomic Science* 10:225-226.
- Stern, C. 1965. Labeling and variety in concept identification with young children. *Journal of Educational Psychology* 56:235-240.
- Tighe, L. S., and Tighe, T. J. 1966. Discrimination learning: Two views in historical perspective. *Psychological Bulletin* 5:353-370.

- Trabasso, T. 1963. Stimulus emphasis and all-or-none learning in concept identification. *Journal of Experimental Psychology* 65:398-406.
- Trabasso, T. 1968. Pay attention. *Psychology Today* 2:30-36.
- Trabasso, T. and Bower, G. H. 1968. *Attention in learning*. New York: Wiley.
- Turnure, C., and Wallach, L. 1965. The influence of contextual variation on the differentiation of parts from wholes. *American Journal of Psychology* 78:481-485.
- White, S. H. 1965. Discrimination learning with ever-changing positive and negative cues. *Journal of Experimental Child Psychology* 2:154-162.
- White, S. H. 1966. Age differences in reaction to stimulus variation. In *Experience, structure, and adaptability*, O. J. Harvey. ed. New York: Springer.

continued from page 338

- Meighan, Thomas. *An Investigation of the Self-concept of Blind and Visually Handicapped Adolescents*. New York: American Foundation for the Blind, 1971. Pp. xi, 43. Paperback, \$2.50.
- Metzner, Ralph. *Maps of Consciousness*. New York: Macmillan, 1971. Pp. 160, \$7.95; paperback, \$3.95.
- Menninger, Karl A. *The Human Mind*. 3d ed. Westminster, Md.: Random House, 1971. Pp. xvii, 517. \$10.00.
- Mikulas, William L. *Behavior Modification: An Overview*. New York: Harper & Row, 1972. Pp. x, 179.
- Miller, Arthur, G., ed. *The Social Psychology of Psychological Research*. New York: The Free Press, 1972. Pp. ix, 454. \$9.95.
- Millon, Theodore, and Diesenhau, H. I. *Research Methods in Psychopathology*. New York: Wiley, 1971. Pp. xv, 191. Paperback.
- Mindess, Harvey. *Laughter and Liberation*. Los Angeles: Nash Publishing, 1971.
- Myklebust, Helmer R., ed. *Progress in Learning Disabilities*. Vol. 2. New York: Grune & Stratton, 1971. Pp. ix, 404. \$13.75.
- Newman, Mildred; Berkowitz, Bernard; and Owen, Jean. *How to Be Your Own Best Friend*. Brooklyn: Lark Publishing, 1971. Pp. 49. \$4.95.
- Osborne, Richard H., ed. *The Biological and Social Meaning of Race*. San Francisco: W. H. Freeman, 1971. Pp. viii, 182. \$6.00; paperback, \$2.95.
- Paivio, Allan. *Imagery and Verbal Processes*. New York: Holt, Rinehart and Winston, 1971. Pp. xi, 596.
- Patterson, Gerald R. *Families: Applications of Social Learning to Family Life*. Champaign, Ill.: Research Press, 1971. Pp. 143. Paperback.
- Pepinsky, Harold B., and Patton, Michael, J., eds. *The Psychological Experiment: A Practical Accomplishment*. Elmsford, N. Y.: Pergamon Press, 1971. Pp. 193. \$8.75.
- Pinkney, Alphonso. *The American Way of Violence*. New York: Random House, 1972. Pp. xiv, 235. \$7.95; paperback, \$1.95.
- Peterson, Severin. *A Unique New Reference Book: A Catalog of the Ways People Grow*. New York: Ballantine Books, 1971. Pp. 368. \$1.65.
- Rabin, A. I. *Kibbutz Studies: A Digest of Books and Articles on the Kibbutz*. East Lansing: Michigan State University Press, 1971. Pp. viii, 124. \$7.50.
- Raknes, Ola. *Wilhelm Reich and Orgonomy*. Baltimore: Penguin Books, 1971. Pp. 183. Paperback, \$1.45.
- Ross, Dorothy. G. *Stanley Hall: The Psychologist as Prophet*. Chicago: University of Chicago Press, 1972. Pp. xix, 482. \$12.75.
- Rubenstein, Richard L. *My Brother Paul*. New York: Harper & Row, 1972. Pp. x, 209. \$5.95.
- Schein, Edgar H. *Organizational Psychology*. 2d ed. Englewood Cliffs, N. J.: Prentice-Hall, 1972. Pp. xiv, 138. \$5.95; paperback, \$2.95.

continued on page 376

FIGURAL AFTEREFFECTS AS OPTICAL ILLUSIONS?

Michael Stadler

University of Münster, Germany

Ganz's theory of figural aftereffects is criticized on these grounds: that the 'attraction effect' he described is an artifact of figural dynamics during the inspection period; that aftereffects cannot be optical illusions between test figure and the afterimage of inspection figure because aftereffects are found without such afterimages; and that lateral inhibition cannot be the underlying mechanism because aftereffects occur in areas outside its field of efficacy.

In recent years Ganz has presented a series of findings (Ganz, 1964, 1966a, 1966b; Ganz and Day, 1965) in support of an entirely new theory of figural aftereffects. He rejected the classic explanations of figural aftereffects, the so-called satiation or fatigue theories of Köhler and Wallach (1944) and Osgood and Heyer (1952), largely because they failed to explain the 'attraction effect' he had generally found in his experiments.

The theory Ganz proposed instead was based on two presuppositions. The first was that the figural aftereffects which result from successive presentation of inspection and test figures are simultaneous *optical illusions* between the test figure and the afterimage of the inspection figure. His second assumption was that the already known repulsion effect and the 'attraction effect' he had found are both explainable on a neurophysiological basis, as the result of the *lateral inhibition* of the contour of the test figure by the contour of the inspection figure—by the state of light or dark adaptation of the retina that the inspection figure causes. The spreading of the effects over an area around the contours can thus be explained by the eye-movement distribution (see Osgood and Heyer, 1952).

In this form, Ganz's theory of figural aftereffects is, on the one hand, more comprehensive than any other theoretical approach, because it sets out to include the optical illusions, and on the other hand, more specific, because it refers explicitly to a well-established neurophysiological mechanism that has been successfully used as an explanation of other groups of phenomena (Ratcliff, 1965).

Though Ganz's theory seemed at first glance to explain most of the phenomena occurring with figural aftereffects, there soon appeared a need for modification of the theory to account for special observations (Pollack, 1967; Immergluck, 1968; Farnè, 1970). In the article at hand, however, the approach to Ganz's theory is of a more fundamental nature: first, his interpretation of the 'attraction effect' will be considered in light of new experimental findings, and second, the two presuppositions of his theory will be subjected to experimental examination, using in most cases the size of closed figures to illustrate the figural aftereffects.

THE 'ATTRACTION EFFECT'

The displacement or repulsion effect described by Köhler and Wallach (1944) is a repulsion of the contour of the test figure by the contour of the inspection figure that surrounds it. According to the so-called distance paradox, this repulsion (an apparent shrinkage of the test figure if the stimuli are concentric circles) should be less very close to the inspection contour and should increase until a certain interfigural distance is reached, beyond which the magnitude of displacement is expected to decrease again.

On the whole, this is what Ganz and other investigators have found. Figure 1 graphs this general relationship in the lower right quadrant, showing how the shrinkage of the test circle is less when the inspection and test figures' contours are very close (at the broken vertical line, they are identical) and then increases until a certain interfigural distance is reached, when the shrinkage of the test circle then begins to decrease.

But Ganz also found that at a very small interfigural distance, of 0–10 min of arc of visual angle between the square and the line he was using as inspection and test figures, the amount of displacement became 'negative' and there was an 'attraction' instead of a repulsion of the contours of inspection and test figures. This he interpreted as an instance of the finding, known since the work of Köhler and Wallach (1944) and illustrated at the intersection of the broken vertical and horizontal lines in Figure 1, that the test circle appears to shrink with congruent inspection and test circles. When these figures are closed, wrote Ganz, this too is an 'attraction effect,' and he cited Oyama (1954) and Yoshida (1952, 1953), who had found similar results using closed circles (see Ganz, 1964; Ganz and Day, 1965, p. 359).

In the opinion of Ganz and Day, the 'attraction effect' was extremely

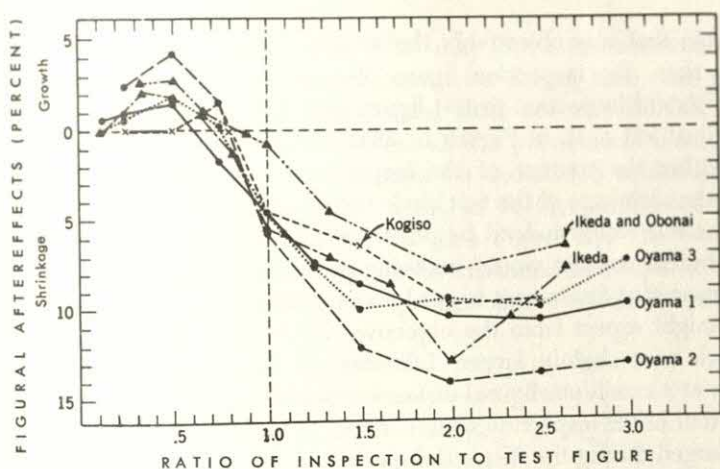


Fig. 1. Figural aftereffects as a function of the ratio of inspection-figure size (from .5 through 3.0) to test-figure size (1), the ratio also serving here as a measure of interfigural distance; the intersection of the broken vertical and horizontal lines indicates physical equality of size of congruent inspection and test circles, shrinkage to the right of the vertical line showing the repulsion effect and shrinkage at the line and to its left showing what Ganz termed the 'attraction effect' (illustration after Sagara and Oyama, 1957; Graham, 1965, p. 555)

damaging to the satiation or fatigue theories because it strikes at their core assumptions that the displacement of contours is away from satiated or fatigued areas of retinal stimulation. And actually, the satiation or fatigue theories could not account for the 'attraction effect' Ganz found, nor do they seem to explain the apparent shrinkage of a test circle the same size as the inspection circle. However, before those theories are declared false, it must be asked whether Ganz interpreted his experimental findings in an appropriate manner or whether his findings could be mere artifacts—specifically, artifacts of considering the physical size and position of the inspection figure as identical with its phenomenal size and position, as Ganz did.¹

In contrast to Ganz's 'attraction' of the test figure by the inspection figure's contours, it seems plausible to think of a phenomenal alteration of the inspection figure during its inspection. Closed inspection figures like circles or squares might thus suffer a phenomenal enlargement, just as curved lines are perceived to be less curved after some time of inspection (Gibson, 1933); such findings are predicted by Köhler's insatiation theory, as the result of 'self-satiation.' If this phenomenal enlargement of, say, a circle did take place during inspection, then a

test circle that was objectively the same size as, or even a little bit greater than, the inspection figure preceding it (that is, when the ratio of size of inspection to test figure was that at the broken vertical line, or just left of it, in Figure 1), that test circle could be seen as lying within the contour of the inspection figure and the repulsion effect (the shrinkage of the test circle) would be expected. In this way, the test circle would indeed be phenomenally diminished, as Figure 1 shows, but its contour would be being pushed *away from* the contour of the inspection figure—not toward it as Ganz's interpretation held and as one might expect from the objective size of both circles when the test circle was slightly larger. Likewise, in Ganz's experiments, the test line at a small interfigural distance would have been falling within the contour of the inspection square if that square had been phenomenally enlarged during the inspection period.

Thus, our first hypothesis was that closed figures suffer a phenomenal enlargement during the inspection period and that the amount of this enlargement is of the same size as the range in which contour 'attraction' has been found, specifically, between 0 and 10 min of arc in interfigural distance. If this hypothesis were confirmed, the 'attraction effect' could be explained as a misinterpretation of experimental findings—explained, that is, in terms not of the *physical* sizes of inspection and test figures but of the influence of the *phenomenal* size of the inspection figure on the phenomenal size of the test figure. This hypothesis of phenomenal enlargement as a causal variable is not part of the satiation or fatigue theories, but it seems necessary to describe the effects that have been obtained. Ganz himself implicitly recognized this need when he remarked that "under both normal and stabilized viewing conditions, it was sometimes observed that the test line appeared within the confines of the afterimage, although ostensibly it was placed by the experimenter at a small interfigural distance outside that figure" (1966b, p. 157).

Experiments to prove the hypothesis of phenomenal enlargement during the inspection period were carried out and have already been published (Erke, 1967; Erke and Stadler, 1967; Stadler and Erke, 1968). In these studies, the inspection circles were shown for both short (50–500 msec) and longer (up to 3 min) inspection periods. The subjects judged the apparent size of the inspection circle against a comparison circle exposed for 50 msec on the other side of the fixation point. An illustration of the results for one subject with a short exposure time is given in Figure 2. For all subjects, the phenomenal enlargement of the inspection figure amounted to 5.1 to 11.7 min of

arc during the first 500 msec. This may be a demonstration of a phenomenon that has been well known for a long time: the gamma movement (Kenkel, 1913; Lindemann, 1922; see also Köhler, 1965).

The hypothesis that closed figures suffer a phenomenal enlargement during the inspection period can thus be considered confirmed. This enlargement levels off after 300 msec, as further experiments have shown (Stadler and Erke, 1968; see also Figure 2). The 'attraction effect' Ganz found can consequently be explained as a *repulsion* effect of the *phenomenal* size of the inspection figure on the phenomenal

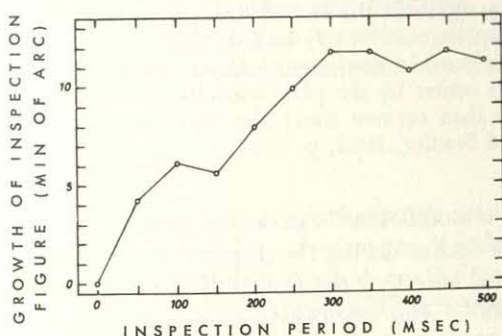


Fig. 2. Phenomenal enlargement of a circle during the inspection period

representation of the test figure, as Figure 3 shows. Thus Ganz's main argument against the satiation or fatigue theories is invalidated.

Indeed, if the repulsion effect did reverse into an 'attraction' effect in the area very close to both sides of an inspection figure, as Ganz supposes, the empirical curves in Figure 1 ought to have a different shape. But the shape of those curves shows that his 'attraction effect' is taking place only on the *outside* of the inspection figure, namely, when the test circle is seen to shrink when it is objectively *larger* than the inspection circle (just to the left of the broken vertical line). When the test circle is seen to shrink when it is smaller than the inspection circle (just to the right of the broken vertical line), this is a repulsion effect, of course. Furthermore, even this one-sided 'attraction effect' becomes a repulsion effect by moving the broken vertical line that indicates physical equality of size of inspection and test figures (or equivalently here, indicates no interfigural distance) a few marks to the left in Figure 1, thereby adjusting the ratio of size of inspection to test figure to the ratio of their *perceived* sizes (or perceived inter-

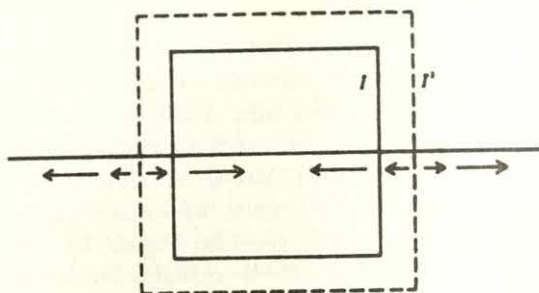


Fig. 3. Schematic interpretation of Ganz's 'attraction effect'; the letter I represents the figure at the beginning of the inspection period, the I' represents the same figure at the end of that period, and the arrows show the direction of displacement. It can be seen that a closed test figure lying at a small distance outside the original closed inspection figure (I) will be displaced to the center by the phenomenally enlarged inspection figure (I'); circles rather than squares could also have been used here (illustration after Erke and Stadler, 1967, p. 327).

figural distance), a modification that seems justified by our findings of an enlargement of figures during the inspection period.

It might be noted here that the results of Ikeda and Obonai (1953) and Duncan (1960), who wanted to ascertain changes in the size of figures with 'self-satiation,' point in the opposite direction: there was an apparent shrinkage of figures during the inspection period. Although these results are contrary to ours, they are not very astonishing because the exposure time of the comparison figures was not controlled. Duncan even quotes an exposure time of 4 sec for those comparison figures, during which time they would probably undergo a change similar to that of the inspection figures. That would produce an entirely different measurement than the ones we cite.

OPTICAL ILLUSIONS AND AFTERIMAGES

The mechanism that bridges the time gap between the presence of the inspection figure and the presence of the test figure in the figural-aftereffect situation is, according to Ganz, the afterimage of the inspection figure. The aftereffect is consequently (Ganz assumes) an optical illusion between the test figure and that afterimage of the inspection figure. Although Ganz's argument does not rest on *reportable* afterimages, his assumption seems to raise several problems.

Length of inspection period

First of all, figural aftereffects have been clearly demonstrated after very short inspection times (Farnè, 1965) that produced afterimages lasting for a very short time only; in one of our experiments such afterimages of inspection figures exposed for less than 500 msec were visible for 100–200 msec at most. Figural aftereffects have also been shown to occur after very long pauses (up to 60 min) between inspection and test periods, during which pauses the afterimages must have soon disappeared (Hetherington, Pick, and Koski, 1964). Afterimages do, we might note, last over 30 min with the method of stabilized retinal images (Barlow and Sparrock, 1964), but those afterimages must be stronger than the ones produced with normal nonstabilized fixation, as in the experiments by Farnè and Hetherington et al.

Intensity of inspection figure

The second difficulty with Ganz's afterimage assumption is that if figural aftereffects are induced by afterimages, it follows that the magnitude of the aftereffects should depend on the intensity of the inspection figures. Ganz (1966b) does quote a series of authors who have shown that inducing figures of high contrast result in figural aftereffects of greater magnitude (Nozawa, 1953; Yoshida, 1955, 1960; Sagara and Oyama, 1957; Pollack, 1958; Graham, 1961). But this problem has not yet been settled, since other authors have not been able to ascertain any noticeable difference of magnitude between the figural aftereffects produced by differences in contrast (Walthall, 1946; Freeburne and Hamilton, 1949; Oyama, 1960; see also Malhotra, 1966).

'Interpretation' of inspection figure

Third, how can Ganz's theory account for the fact that figural aftereffects are produced by perceptual structures that leave no afterimages—for the fact that the effects of attitude and of figure interpretation, for example, cause considerable differences in magnitude of the figural aftereffects, as has been shown by Story (1959), Weitzman (1963, 1966), Willems (1967), and Erke and Schulte (1968) under different experimental arrangements? A similar instance from another recent experiment (Stadler, 1969) is shown in Figure 4.

In the left half of Figure 4, the two test parallels are shown on the space between the two preceding inspection figures; in the right half

of Figure 4, they are shown at the center of the area that had been occupied by what seemed to be a single inspection figure (a vase on its side) whose middle was masked by a screen. In the latter case, we

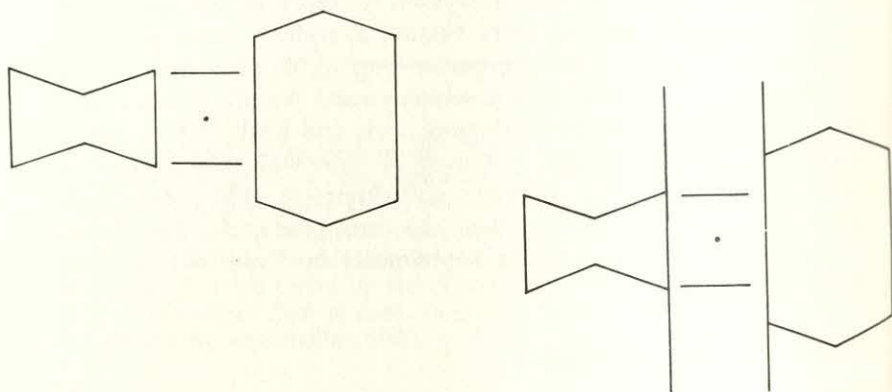


Fig. 4. Control (left panel) and experimental (right panel) displays producing figural aftereffects without afterimages but with amodally completed contours (from Stadler, 1969, p. 28)

found, the test parallels were affected by the so-called *amodally completed contours*, the contours present whenever subjects are describing complete structures even though the corresponding physical (distal stimulus) conditions show incomplete structures (Michotte and Burke, 1951; Michotte, Thinès, and Crabbé, 1966). The figural aftereffect with the experimental display shown in the right panel of Figure 4 accorded with the Ponzo illusion, in that the test parallels were perceived to converge to the right. This aftereffect was greater by .35 mm, a very significant difference, than that with the display shown in the left panel of Figure 4.

An even clearer instance, this time of visible (not masked) contours without any counterpart in the physical stimulus condition, is represented by the so-called *virtual contours* (Kanizsa, 1955). In Figure 5, it can easily be seen that the slight, but sharply separated, apparent contours between the black disks do not produce an after-image when one fixates the whole figure for a while. Farnè (1968) and Stadler and Dieker (1969, 1972) proved independently that aftereffects are induced by such virtual contours, a result that cannot

be reduced to an effect of the physical (distal stimulus) condition. Farnè (1970) has already pointed out the difficulties that this finding poses for Ganz's theory.

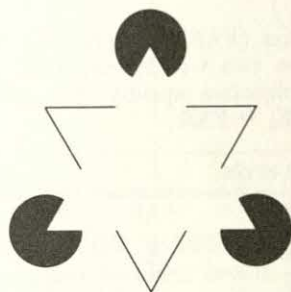


Fig. 5. Display producing figural aftereffects without afterimages but with virtual contours (from Kanizsa, 1955, p. 16)

Visibility of afterimage

There is yet a fourth difficulty with Ganz's theory. Until now, Ganz could eventually answer the objections outlined above by saying that his theory required no visible afterimages to evoke the figural aftereffects—only a process of light adaptation on *some* level of the visual system. He assumed, in other words, that subliminal differences in brightness suffice in themselves for figural aftereffects. The evidence, however, suggests that those differences in brightness must be represented phenomenally (must be above threshold). As Hochberg, Day, and Hardy (1960) observed, figural aftereffects seem to occur only if the inspection contour is—if just barely—noticeable. It seemed worthwhile to study this problem a bit more closely. To begin with, we carried out two preliminary experiments.

In the first, the hypothesis that figural aftereffects can be induced by *subliminal* inspection figures was tested. The inspection figure was a photograph of a gray circle, 4 cm in diameter, on a gray background. Contrast between circle and ground was so low that the circle could not be noticed during a 5-sec tachistoscopic exposure. For the seven subjects tested, this subliminal inspection figure did not produce any significant figural aftereffects on test circles that were greater (5 cm) or smaller (3 cm) than it; see Table 1. We must concede that it was not possible to prove whether our stimuli did affect the visual system at all. For that reason, the outcome of this experiment is difficult

to interpret; but in connection with the outcomes of the two next experiments, which were performed under the same general conditions, it is suggestive.

Table 1. Figural aftereffects (FAE, shown in mm) induced by a subliminal 4-cm inspection circle on two test figures; PSE_I and PSE_0 respectively represent the points of subjective equality with and without the inspection figure, so that $PSE_I - PSE_0 = FAE$

Subject	5-cm test circle			3-cm test circle		
	PSE_I	PSE_0	FAE	PSE_I	PSE_0	FAE
A	48.6	48.4	.2	28.3	28.6	-.3
B	51.3	51.3	.0	29.8	29.7	.1
C	49.0	48.5	.5	29.1	29.8	-.7
D	49.0	49.1	-.1	28.8	29.3	-.5
E	49.6	50.7	-1.1	29.3	29.3	.0
F	49.8	49.8	.0	29.2	29.0	.2
G	50.3	51.1	-.8	29.4	28.8	.6
Mean			-.2			-.1
<i>t</i> test			$p > .10$			$p > .50$

Note: Both test circles were viewed for 100 msec; the inspection figure was presented for 5 sec. Positive numbers show expected shrinkage of the test circle; negative numbers show a growth of the test circle.

In the second preliminary experiment, the hypothesis that figural aftereffects are induced by 'inspection' figures that are only *imagined* was tested. A white surface with a fixation point in the middle (see the dot in Figure 6) was exposed tachistoscopically to 11 subjects who were instructed to hold strong fixation and to imagine a circle the size of a five-mark piece (2.9 cm in diameter) to the left of the fixation point (at the small cross in Figure 6). Centered on exactly the same cross one of the two test circles was then exposed after the 5-sec inspection period, so that if the subjects had projected their imagined circle in the right place, the test circle must have been surrounded by its imagined contour. The subjects were not allowed to see a five-mark piece before or during the experiment, even if they wished to. This experimental condition was preceded and followed by control conditions exactly like it but without the instruction to imagine the circle. The mean of the values of these two control conditions is plotted in Table 2 under PSE_0 . The outcome shows significant figural aftereffects (shrinkage of the test circle) in the experimental condition.

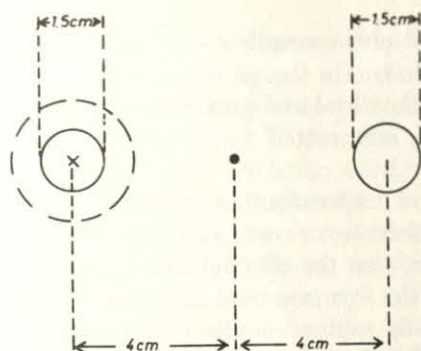


Fig. 6. Experimental arrangement producing figural aftereffects without afterimages but with imagined contours; broken circle is where imagined figure was to be projected, solid circles are the test figures (one for comparison)

Table 2. Figural aftereffects (FAE, shown in mm) induced by an imagined 2.9-cm inspection circle on a test figure; PSE_I and PSE_0 respectively represent the points of subjective equality with and without the inspection figure, so that $PSE_I - PSE_0 = FAE$

Subject	1.5-cm test circle		
	PSE_0	PSE_I	FAE
A	15.1	12.7	2.4
B	15.8	14.5	1.3
C	15.3	14.7	.6
D	15.0	15.7	-.7
E	14.6	14.2	.4
F	15.5	15.4	.1
G	14.9	14.5	.4
H	14.5	14.3	.2
I	16.9	16.4	.5
K	16.3	14.4	1.9
L	16.2	15.4	.8
Mean			.7
<i>t</i> test (one-tailed)			$p < .01$

Note: Each test figure was viewed for 100 msec; the inspection figure was imagined for 3 sec. Positive numbers show shrinkage of the test circle, the expected figural aftereffect; negative numbers show a growth of the test circle.

The two experiments just briefly outlined show, first of all, that inspection figures which were given in the stimulus condition but

were not presented phenomenally (were subliminal) did not induce any figural aftereffects. On the other hand, 'inspection' figures which were not given in the distal and proximal stimulus condition but were only phenomenally represented (were imagined) did induce figural aftereffects.

This led us to an experiment in which the transition from non-perceived to perceived figures was examined, by the method of constant stimuli, to see what the effect of that transition might be on any figural aftereffects the figures produced (Stadler and Müskens, 1972). We tachistoscopically exposed a circle with a 4-cm diameter for inspection periods of 1, 2, 4, 8, 16, 32, 64, 128, and 256 msec. Without pause, this inspection circle was followed by one test circle with a 3-cm diameter and one comparison circle with approximately a 3-cm diameter, the seven steps (from 27 to 33 mm) of variation in its size being presented six times each in random order. The nine subjects did not know whether an inspection figure was to be expected or not on any given trial. Then, after each individual trial, each subject was asked whether or not he had perceived the inspection figure. Figure 7 shows the percentage of subjects reporting an aftereffect (here, shrinkage) and the percentage of individual trials on which they reported that they had seen the inspection figure. The two resulting curves obviously agree well with one another.

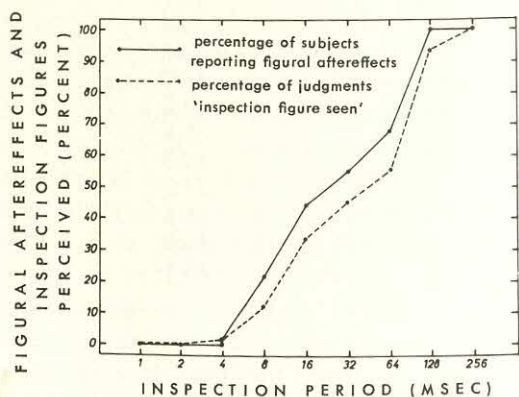


Fig. 7. Figural aftereffects and perception of the inspection figure as functions of the length of the inspection period

In words, Figure 7 shows that as long as no inspection figure was perceived, there was no figural aftereffect, that is, no perceived shrink-

age of the test figure; not until the percentage of inspection figures that were perceived had increased did figural aftereffects differing from null appear; and when the percentage of inspection figures perceived (as reported by all subjects) rose to 100%, then all of the subjects reported figural aftereffects. The correlation coefficients between the two variables, magnitude of figural aftereffect and percentage of judgments 'inspection figure seen,' were calculated independently for the inspection periods of 8, 16, 32, 64, and 128 msec. (At the inspection periods of 1, 2, 4, and 256 msec, the two variables did not differ between subjects; we found only values of 0% or 100% for 'inspection figure seen.') The resulting five correlations of .73, .85, .60, .94, and .33 were statistically homogeneous ($\chi^2 = 6.9, df = 4$), so it can be concluded that the two phenomenal variables are highly related independently of the stimulus variable of length of inspection period. It can therefore be said that figural aftereffects occur to the same extent that the inspection figure is (phenomenally) perceived by the subjects. This result is in accordance with the findings that virtual contours, amodally completed contours, and even imagined contours do produce significant figural aftereffects, even though those contours are only phenomenally represented (are not given in the objective stimulus condition) and even though they do not produce any afterimage (apart from the solitary findings of Downey, 1901).

In sum, our experiments showed that figural aftereffects are strictly bound to phenomenally represented inspection contours. Subliminal contours will not produce the aftereffects, but imagined contours will.

Locus of stimulation by inspection figure

It would be extremely damaging for Ganz's theory if it were found that the occurrence of figural aftereffects is determined not by the locus of retinal stimulation, the point upon which the inspection figure is projected, but by the position of that figure in phenomenal space. This is so because it is well known that afterimages are stringently restricted to the retinal area stimulated by the inducing stimulus.

Our experimental arrangement was to expose an inspection circle (I; top panel, Figure 8) in the middle of the tachistoscopic field; to the right, at the dot, was the fixation point. A 5-sec inspection period was followed by a 500-msec pause. The subjects were instructed then (middle panel, Figure 8) to turn their eyes to fixate another fixation point left of the middle, and the first couple of test circles (T_1 and T_2) was exposed, T_1 at the same retinal position and T_2 at the same

egocentric position as the inspection figure. These test figures remained visible for 500 msec until a new pause of 500 msec, during which (bottom panel, Figure 8) the subjects turned their eyes back to the right where the fixation point of the second couple of test circles (T_3 and T_4) was located, these test figures also being visible

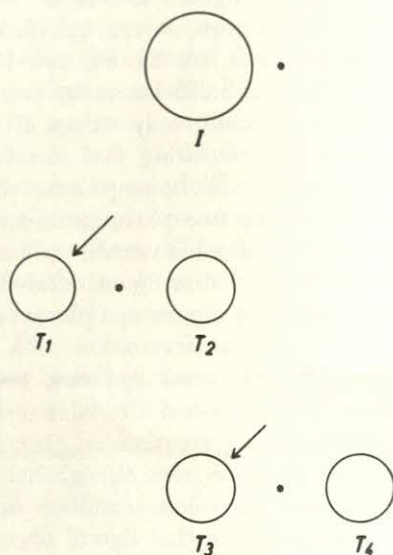


Fig. 8. Experimental arrangement producing figural aftereffects independent of locus of retinal stimulation; the inspection figure (I) and two pairs of test circles (T_1 and T_2 ; T_3 and T_4) were shown successively with change in fixation point (small dots); the arrows show which circle fell into the retinal area previously affected by the inspection circle

for 500 msec. After this sequence, the subjects made their judgments for the first and the second couple of test couples successively.

The results showed no effect of whether the subjects judged the right test circles (T_2 , T_4) with regard to the left (T_1 , T_3) or judged the outside circles (T_1 , T_4) with regard to the inside circles (T_2 , T_3), and the data from the two groups were averaged (PSE_1). Also, control conditions without the inspection figure were carried out (PSE_0). Table 3 shows the outcome for our ten subjects. The T_2 circle was perceived as reduced about .28 mm on the average, and T_3 as reduced about .18 mm on the average. The lesser figural aftereffect at T_3 can be explained as an effect of the longer pause (1.5 sec) between the

end of the inspection period and this particular judgment.

Table 3. Figural aftereffects (FAE, shown in mm) dependent on phenomenal location of an inspection circle for two of four test figures; PSE_1 and PSE_0 respectively represent the points of subjective equality with and without the inspection figure, so that $PSE_1 - PSE_0 = FAE$

Subject	Test circles T_1 and T_2 FAE found for T_2			Test circles T_3 and T_4 FAE found for T_3		
	PSE_0	PSE_1	FAE	PSE_0	PSE_1	FAE
A	25.83	25.17	.66	24.75	24.33	.42
B	24.83	24.75	.08	25.25	25.42	-.17
C	25.25	25.00	.25	24.75	24.83	-.08
D	25.25	25.33	-.08	24.92	24.75	.17
E	24.75	24.33	.42	25.67	25.17	.50
F	24.00	23.75	.25	25.75	25.75	.00
G	24.83	24.75	.08	24.83	24.33	.50
H	25.58	25.00	.58	25.33	25.50	-.17
I	24.42	23.83	.59	25.42	25.17	.25
K	24.58	24.58	.00	25.00	24.67	.33
Mean			.28			.18
<i>t</i> test (one-tailed)			$p < .01$			$p < .05$

Note: Each pair of test circles (T_1 and T_2 ; T_3 and T_4) was viewed for 500 msec; the inspection figure was exposed for 5 sec. Positive numbers show expected shrinkage of test figure; negative numbers show its enlargement.

The very important result is the first demonstration that only the test figure that has the same locus in phenomenal space (egocentric localization) as the inspection figure suffers figural aftereffects induced by that inspection figure. If retinal coordinates alone were relevant for figural aftereffects, then both of the test figures in Figure 8 that are marked with arrows (T_1 and T_3) should have endured figural aftereffects—should have diminished. In other words, it is the absolute (egocentric) and not the relative (retinal) localization that is decisive for figural aftereffects.² At T_3 the egocentric and the retinal localization are equivalent, as they have been in all previous experiments; at T_1 and T_2 the egocentric and the retinal localization are, however, different according to the preceding change of fixation.

In sum, this experiment showed that under certain conditions figural aftereffects occur in places where no afterimage of the inspection figure is situated. For a more detailed description of this experimental arrangement and further results, see Crabus and Stadler, 1971.

LATERAL INHIBITION

Ganz's explanation of figural aftereffects as a function of lateral inhibition (see the discussions by Day, 1962, and Deutsch, 1964) seems very elementary and plausible if one grants that those aftereffects are induced by the afterimage of the inspection contours. That is to say, there must be two adjacent simultaneous neural processes to make lateral inhibition at all effective. This explanation varies from those of Köhler and Wallach (1944) and Osgood and Heyer (1952), who introduced the concepts of satiation or fatigue to explain how it is that the effectiveness of the inspection figure's contour outlasts its presentation.

The mechanism of lateral inhibition as based on the organization of the receptive fields of the visual system can be regarded as very well corroborated (Hartline, 1940; Kuffler, 1953; Hubel and Wiesel, 1959; Ratliff, 1965). The importance of this mechanism as a physiological correlate of perceptual phenomena has been sufficiently demonstrated (Jung, 1961; Spillmann, 1963).

It is, however, questionable whether lateral inhibition in the first postretinal layers can be taken as the *only* physiological basis for the explanation of perceptual phenomena as complex as figural aftereffects or optical illusions. This is so because lateral inhibition is a form of stimulus processing whose functional entities extend over relatively small areas in the nervous system. Indeed, as Ganz himself notes (1966a), with reference to an investigation of Leibowitz, Mote, and Thurlow (1953), "most of the [so-called] inhibition occurs within the initial 10 min of arc [interfigural distance on the fovea] whether the inducing stimulus is strong or weak." But according to Baumgartner (1960), the receptive fields—the area in which lateral inhibition can occur—have a diameter of .025 mm in the foveal region of the human retina. This corresponds to only 5 min of arc there on the fovea, a more limited range of efficacy of lateral inhibition than that which Ganz posits.

We designed an experiment to test Ganz's hypothesis that figural aftereffects depend on lateral inhibition. Our method was by means of variation of sizes of inspection figures in the periphery of the retina, where, according to Kornhuber and Spillmann (1964), the receptive fields for lateral inhibition are somewhat greater, extending to 1–2 deg of arc. Three different sizes of inspection figure were used: 3 cm, which subtended 72 min of arc there on the periphery; 9 cm, subtending 3 deg 36 min; and 15 cm, subtending 6 deg of arc.

The ratios of inspection to test figure were 1:2, 1:1.5, 1:1, 1:.75, and 1:1.5, so that the size of the smallest (1.5-cm) test circle was 36 min of arc in diameter, that of the greatest (30-cm) test circle 12 deg in diameter. The ratios indicate the figural aftereffects that would be expected: with the inspection circle larger than the test circle (1:75 and 1:1.5), the aftereffect would be shrinkage of the test circle; with the test circle larger than the inspection circle (1:2 and 1:1.5), the aftereffect would be growth of the test circle; with the inspection and test circles the same size (1:1), we would have the special case discussed earlier in this paper.

The thickness of the line was 1 mm for all figures, and all were shown on a translucent screen by means of two diapositive projectors from a distance of 190 cm. Each of the seven subjects sat on the other side of the screen from the projectors, 145 cm away from it. With the fixation point in the middle of the screen, the inspection and test circles were exposed in the periphery (at between 20 and 60 deg from the line of vision). The interfigural distance between inspection and test contours themselves ranged from 9 min of arc for the 3-cm (72 min) inspection circle with a 2.25-cm (54 min) test circle, to 3 deg for the 15-cm (6 deg) inspection circle with a 30-cm (12 deg) test circle.

Thus, the variation of inspection-figure and test-figure size covered and extended beyond the range of interfigural distance for which, according to Kornhuber and Spillmann (1964), there are receptive fields for lateral inhibition in the retinal periphery, namely, the range of between 1 and 2 deg of visual angle. Beyond this range, the effect of lateral inhibition in the postretinal layers should be limited and thus, by Ganz's theory, produce a reduction or a complete disappearance of figural aftereffects.

Figure 9 shows our results. For all interfigural distances, the aftereffects were as expected (whether shrinkage or growth of the inspection circle, depending on the ratio of inspection-figure to test-figure size), namely, the smaller aftereffects with smaller interfigural distances as postulated by the distance paradox. There was no significant decrease in magnitude, nor any disappearance, of the growth aftereffects in the range between 1 and 2 deg of visual angle—nor, contrary to the prediction from Ganz's theory, beyond 2 deg. Interestingly also, the effect (here, shrinkage) was least when inspection and test figures were the same size; that is, when the ratio of their sizes was 1:1, which is, as Köhler and Wallach (1944) stressed, a special condition in the figural-aftereffect experiment, a condition dis-

cussed at some length in the section above on the 'attraction effect.'

In sum, our experiment showed that the figural aftereffects (growth)

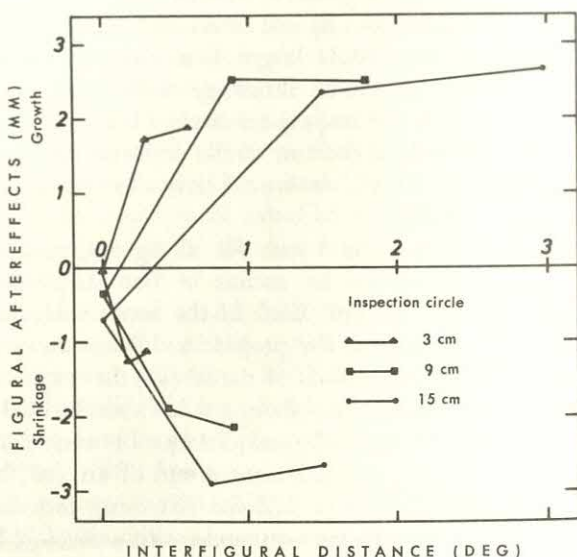


Fig. 9. Figural aftereffects as a function of interfigural distance (in degrees of visual angle) between inspection figure and test figure; the five points on each of the three curves (representing three sizes of inspection circle) show, beginning at the upper right of each curve, the values of the figural aftereffects for these ratios of inspection-figure size to test-figure size: 1:2, 1:1.5, 1:1, 1:0.75, and 1:0.5.

did not diminish or completely disappear in or beyond the range of 1-2 deg of visual angle of interfigural distance. Beyond that range, where Ganz's theory would predict the effect of lateral inhibition and thus the aftereffects to diminish, we found instead figural aftereffects of constant magnitude.

CONCLUSIONS

We were able to show, with the help of some new experiments, that Ganz's main objection to the satiation or fatigue theories of figural aftereffects does not, after all, apply. Instead, the 'attraction effect' he described (Ganz and Day, 1965) can be explained under the special

conditions of their experiment as an artifact of figural dynamics during the inspection period.

We then tested the basic assumptions of Ganz's own theory. It was found that figural aftereffects do not exclusively result from inspection figures that produce afterimages but also occur with inspection figures that arise on a very high level in the visual system and have no counterpart in the proximal or distal stimulus. Indeed, so-called virtual contours, amodally completed contours (according to Metzger, 1966, a paradox of the physiology of stimulation), and even imagined 'inspection' figures evoked aftereffects on their successive test figures. Thus, the hypothesis that figural aftereffects are simultaneous optical illusions of the test figure with the afterimage of the inspection figure can no longer be supported. Moreover, it seems that the aftereffects occur only when an inspection figure is consciously perceived; a subliminal inspection figure does not induce them.

It was also found that the occurrence of figural aftereffects is not necessarily bound to the retinal locus of the inspection figure. On the contrary, those aftereffects occurred where the inspection figure was located in phenomenal space independent of its position on the retina. In other words, egocentric localization, and not retinal localization, is the decisive factor. This fact has been neglected in previous experiments, perhaps because in the usual situation egocentric and retinal localization are identical. Only by means of a change of fixation in the pause between inspection and test periods can the difference between egocentric and retinal localization be demonstrated.

These experiments are evidence not only against Ganz's theory but also against all theories which suppose that figural aftereffects start from the retina; the theory of Motokawa, Nakagawa, and Kohota (1957), which is based on the mechanism of retinal induction, is but one example. Instead, our experiments suggest that the process underlying figural aftereffects takes place on a level of the visual system where eye movements have already been compensated for, so that phenomenal constancy of direction is present. This suggestion gains support from the findings that figural aftereffects are induced by virtual contours (as well as by 'real' contours given in the distal and proximal condition), since virtual contours do not arise from stimuli on the retina but are the result of the processing of the stimulus information in the visual system.

The experiments described in this paper do not tell us anything about the connection between optical illusions and figural after-

effects, however. Many discussions can be found in the literature on this topic (Ikeda and Obonai, 1955; Rausch, 1966; Adam, 1966; Dreistadt, 1968). Some authors (Robinson, 1968; Wagner, 1968) hold that Ganz's theory accounts only for certain 'distortions' such as the Poggendorf, Zöllner, and Orbison figures, whereas figural after-effects are more akin to some phenomena of the other group, the 'illusions proper' such as the Ponzo and Müller-Lyer figures.

Thus, the main question with which Ganz's theory attempted to deal—the question of the connection between figural aftereffects and optical illusions—still presents many unsolved problems. But since Ganz's attempt to bridge the gap between inspection and test period with the help of the afterimage of the inspection figure seems to us no adequate solution, we hold that it is justified to continue using 'satiation' as a construct, delaying its replacement by a specific physiological process or mechanism (say, the 'field processes' of Köhler) until more physiological evidence can be quoted.

Indeed, it might even turn out that the 'satiation process' is not the direct physiological basis for figural aftereffects but rather triggers a specific interaction with the processing of the test contours when that test figure appears, an interaction in which lateral inhibition might be important (Stadler, 1968). If that were so, the difference between optical illusions and figural aftereffects might then be that with illusions there is a direct interaction of the processing of contours, whereas with aftereffects the processing of the contours of the inspection figure can influence the processing of the contours of the test figure only by interposition of a satiation pattern—a one-sided directive effect.

But if lateral inhibition plays an important role in the formation of optical illusions and figural aftereffects, then this formation cannot occur on the retinal level. As our last experiment showed, the occurrence of figural aftereffects is not limited to the area of efficacy of lateral inhibition, namely, to the range of 1–2 deg on the retinal periphery. On the contrary, the magnitude of figural aftereffects remained constant with increasing interfigural distance beyond this range. Yet perhaps on higher levels of the visual system, the operation of lateral inhibition may have another dimension. Hubel and Wiesel (1962) specify the size of cortical receptive fields of the cat as 4–10 deg, the sort of finding that if extended to humans might explain the displacement effect of simultaneous and successive presentation. It would also agree with the conclusion, drawn from the experimental findings cited above, that figural aftereffects can arise only on a higher level of the visual system, not on the retina.

Notes

This study was supported by many collaborators to whom the author gives thanks: to Erika Detering, Ulrich Holtmann, Werner Müskens, and Karin Gudrun Raschen; as well as to Daniel O'Connell and Falk Seeger for their translation into English and much useful advice. Professor Stadler's present address is at the Psychologisches Institut der Universität Münster, Rosenstrasse 9, 44 Münster, Western Germany. Revision received for publication July 28, 1971.

1. The distinction between physical size (shape, position, etc.) and phenomenal size (shape, position, etc.) introduces a difference of a fundamental kind: the difference between the physical object as represented in a theoretical language or operationalized by measurement, and the phenomenal object as represented in the phenomenal world of the individual. Thus, to a single physical or transphenomenal world correspond the always different phenomenal worlds of different individuals (see Brandt and Metzger, 1969; Metzger, 1970). Because of this complexity, three steps of stimulus processing must be differentiated (see Koffka, 1935): the first has to do with the *distal* stimulus, of which an object emitting electromagnetic light waves would be an example; the second step concerns the *proximal* stimulus, such as the distribution of light energy on the retina caused by those light waves; and the third is with the *phenomenal* object, say, the perceived sheet of paper. The distinction between these three successive steps of stimulus processing must be kept in mind to recognize the importance of the findings presented in this paper.

2. Further research has shown that figural aftereffects can be divided into peripheral (retinal) and central (egocentric) effects, which under the usual figural-aftereffect conditions work in the same direction (Crabus and Stadler, 1971; Stadler and Crabus, 1972).

References

- Adam, June. 1966. The relationship between visual illusions and figural aftereffects. *Australian Journal of Psychology* 18:130-136.
- Barlow, H. B., and Sparrock, J. M. B. 1964. The role of afterimages in dark adaptation. *Science* 144:1309-1314.
- Baumgartner, G. 1960. Indirekte Grössenbestimmung der rezeptiven Felder der Retina beim Menschen mittels der Hermannschen Gittertäuschung. *Pflügers Archiv* 272:21-22.
- Brandt, L. W., and Metzger, W. 1969. 'Reality,' what does it mean? *Psychological Reports* 25:127-135.
- Crabus, H., and Stadler, M. 1971. Über Wahrnehmungsprozesse in den Koordinaten der Netzhaut und des anschaulichen Raumes. *Psychologische Forschung* 34:325-342.
- Day, R. H. 1962. Excitatory and inhibitory processes as the basis of contour shift and negative aftereffect. *Psychologia* 5:185-193.
- Deutsch, J. A. 1964. Neurophysiological contrast phenomena and figural aftereffects. *Psychological Review* 71:19-26.
- Downey, June E. 1901. An experiment on getting an after-image from a mental image. *Psychological Review* 8:42-55.

- Dreistadt, R. 1968. The effect of figural aftereffects on geometrical illusions. *Journal of Psychology* 69:63-73.
- Duncan, C. P. 1960. Effect of self-satiation on perceived size of a visual figure. *Journal of Experimental Psychology* 60:130-136.
- Erke, H. 1967. Grössenänderungen optisch wahrgenommener Figuren während der Inspektionszeit. *Psychologische Forschung* 31:63-90.
- Erke, H., and Schulte, D. 1968. Figur-Auffassung und figurale Nachwirkungen. *Psychologische Forschung* 32:1-13.
- Erke, H., and Stadler, M. 1967. Einige räumliche und zeitliche Bedingungen figuraler Nachwirkungen. In *Bericht über den 25. Kongress der Deutschen Gesellschaft für Psychologie*, pp. 324-329.
- Farnè, M. 1965. Figural aftereffects with short exposure time: An interpretation of the so-called "subliminal perception." *Psychologische Forschung* 28:519-534.
- Farnè, M. 1968. Alcune osservazioni con linee virtuali e margini quasi percettivi. *Bollettino della Società Italiana di Biologia Sperimentale* 44:1613-1616.
- Farnè, M. 1970. Some observations on Ganz's theory. *Acta Psychologica (Amsterdam)* 32(2):192-195.
- Freeburne, C. M., and Hamilton, C. E. 1949. The effect of brightness on figural aftereffect. *American Journal of Psychology* 62:567-569.
- Ganz, L. 1964. Lateral inhibition and the location of visual contours: An analysis of figural aftereffects. *Vision Research* 4:465-481.
- Ganz, L. 1966a. Mechanism of the figural aftereffects. *Psychological Review* 73:128-150.
- Ganz, L. 1966b. Is the figural aftereffect an aftereffect? *Psychological Bulletin* 66:151-165.
- Ganz, L., and Day, R. H. 1965. An analysis of the satiation-fatigue mechanism of figural aftereffects. *American Journal of Psychology* 78:345-361.
- Gibson, J. J. 1933. Adaptation, aftereffect, and contrast in the perception of curved lines. *Journal of Experimental Psychology* 16:1-31.
- Graham, C. H., ed. 1965. *Vision and visual perception*. New York: Wiley.
- Graham, E. H. 1961. Figural aftereffects as functions of contrast, area, and luminance of the inspection figure. *Psychologia* 4:201-208.
- Hartline, H. K. 1940. The receptive fields of optic nerve fibers. *American Journal of Physiology* 130:690-699.
- Hetherington, E. M., Pick, H. L., Jr., and Koski, C. H. 1964. Studies of the influence of differential visual experience on figural aftereffects. *Journal of Experimental Psychology* 68:466-469.
- Hochberg, J. E., Day, R. H., and Hardy, D. 1960. Hue- and brightness-differences, contours, and figural aftereffects. *American Journal of Psychology* 75:638-639.
- Hubel, D. H., and Wiesel, T. N. 1959. Receptive fields of single neurons in the cat's striate cortex. *Journal of Physiology* 148:574-591.
- Hubel, D. H., and Wiesel, T. N. 1962. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. *Journal of Physiology* 160:106-154.
- Ikeda, H., and Obonai, T. 1953. The quantitative analysis of figural after-

- effects: II. On "self-satiation." *Japanese Journal of Psychology* 24:179-192.
- Ikeda, H., and Obonai, T. 1955. Studies in figural aftereffects: IV. The contrast-confluence illusion of concentric circles and the figural aftereffect. *Japanese Psychological Research* 2:17-23.
- Immergluck, L. 1968. Further comments on "Is the figural aftereffect an aftereffect?" *Psychological Bulletin* 70:198-200.
- Jung, R. 1961. Neuronal integration in the visual cortex and its significance for visual information. In *Sensory communication*, ed. W. A. Rosenblith, pp. 627-674. Cambridge, Mass.: The M.I.T. Press.
- Kanizsa, G. 1955. Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia* 49:7-30.
- Kenkel, F. 1913. Untersuchungen über den Zusammenhang zwischen Erscheinungsgrösse und Erscheinungsbewegung bei einigen sogenannten optischen Täuschungen. *Zeitschrift für Psychologie* 67:358-449.
- Köhler, W. 1965. Unsolved problems in the field of figural aftereffects. *Psychological Record* 15:63-83.
- Köhler, W., and Wallach, H. 1944. Figural aftereffects: An investigation of visual processes. *Proceedings of the American Philosophical Society* 88:269-357.
- Koffka, K. 1935. *Principles of gestalt psychology*. London: Routledge & Kegan Paul.
- Kornhuber, H. H., and Spillmann, L. 1964. Zur visuellen Feldorganisation beim Menschen: Die rezeptiven Felder im peripheren und zentralen Gesichtsfeld bei Simultankontrast, Flimmerfusion, Scheinbewegung und Blickfolgebewegung. *Pflügers Archiv* 279:5-6.
- Kuffler, S. W. 1953. Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology* 16:37-68.
- Leibowitz, H., Mote, F. A., and Thurlow, W. R. 1953. Simultaneous contrast as a function of separation between test and inducing fields. *Journal of Experimental Psychology* 46:453-456.
- Lindemann, E. 1922. Experimentelle Untersuchungen über das Entstehen und Vergehen von Gestalten. *Psychologische Forschung* 2:5-60.
- Malhotra, M. K. 1966. Figurale Nachwirkungen. *Psychologische Forschung* 30:1-104.
- Metzger, W. 1966. Figural-Wahrnehmung. In *Wahrnehmung und Bewusstsein*, ed. W. Metzger, *Handbuch der Psychologie*, vol. 1/I, pp. 693-744. Göttingen: Hogrefe.
- Metzger, W. 1970. The phenomenal-perceptual field as a central steering mechanism. Paper read at the Banff Conference on Theoretical Psychology.
- Michotte, A., and Burke, L. 1951. Une nouvelle énigme de la psychologie de la perception: Le "donné amodal" dans l'expérience sensorielle. In *Proceedings of the 13th International Congress of Psychology*, pp. 179-180. Stockholm.
- Michotte, A., Thines, G., and Crabbé, Geneviève. 1966. Die amodalen Ergänzungen von Wahrnehmungsstrukturen. In *Wahrnehmung und Bewusstsein*, ed. W. Metzger, *Handbuch der Psychologie*, vol. 1/I, pp. 978-1002. Göttingen: Hogrefe.

- Motokawa, K., Nakagawa, D., and Kohata, T. 1957. Figural aftereffects and retinal induction. *Journal of General Psychology* 57:121-135.
- Nozawa, S. 1953. Prolonged inspection of a figure and the aftereffect thereof. *Japanese Journal of Psychology* 23:217-234, 24:47-58, 85-87.
- Osgood, C. E., and Heyer, A. W., Jr. 1952. A new interpretation of figural aftereffects. *Psychological Review* 59:98-118.
- Oyama, T. 1954. Experimental studies of figural aftereffects: II. Spatial factors. *Japanese Journal of Psychology* 25:195-206, 223-224.
- Oyama, T. 1960. Figural aftereffects as a function of hue and brightness. *Japanese Psychological Research* 2:74-80.
- Pollack, R. H. 1958. Figural aftereffects: Quantitative studies of displacement. *Australian Journal of Psychology* 10:269-277.
- Pollack, R. H. 1967. Comment on "Is the figural aftereffect an aftereffect?" *Psychological Bulletin* 68:59-61.
- Ratliff, F. 1965. *Mach bands: Quantitative studies of neural networks in the retina*. San Francisco: Holden-Day.
- Rausch, E. 1966. Probleme der Metrik (geometrisch-optische Täuschungen). In *Wahrnehmung und Bewusstsein*, ed. W. Metzger, *Handbuch der Psychologie*, vol. 1/I, pp. 776-865. Göttingen: Hogrefe.
- Robinson, J. O. 1968. Retinal inhibition in visual distortion. *British Journal of Psychology* 59:29-36.
- Sagara, M., and Oyama, T. 1957. Experimental studies on figural aftereffects in Japan. *Psychological Bulletin* 54:327-338.
- Spillmann, L. 1963. Zur Feldorganisation der visuellen Wahrnehmung beim Menschen. Unpublished doctoral dissertation, Universität Münster.
- Stadler, M. 1968. Funktionelle Grundlagen der Wahrnehmung von Figuren. Unpublished doctoral dissertation, Universität Münster.
- Stadler, M. 1969. Figurale Nachwirkungen als Methode zur Untersuchung funktionaler Zusammenhänge im Bereich der Figuralwahrnehmung. *Studia Psychologica* 11:23-31.
- Stadler, M., and Crabus, H. 1972. Peripheral vs. central factors in figural aftereffects. Paper read at the 20th International Congress of Psychology, Tokyo.
- Stadler, M., and Dieker, J., 1969. Margini quasi-percettivi e after-effects figurali. *Rivista di Psicologia* 63:95-104.
- Stadler, M., and Dieker, J. 1972. Untersuchungen zum Problem virtueller Konturen in der visuellen Wahrnehmung. *Zeitschrift für experimentelle und angewandte Psychologie* 19:325-350.
- Stadler, M., and Erke, H. 1968. Über einige periodische Vorgänge in der Figuralwahrnehmung. *Vision Research* 8:1081-1092.
- Stadler, M., and Müskens, W. 1972. Ein Indikator für die Bewusstheit visueller Wahrnehmungen. In *Bericht über den 27. Kongress der Deutschen Gesellschaft für Psychologie* (Kiel, 1970). Göttingen: Hogrefe.
- Story, Anne W. 1959. Figural aftereffects as a function of the perceived characteristics of the inspection-figure. *American Journal of Psychology* 72:46-56.
- Wagner, H. L. 1968. The illusions and Ganz's theory of contour displacements. *British Journal of Psychology* 59:361-367.

- Walthall, W. J., Jr. 1946. The Köhler effect. *American Journal of Psychology* 59:152-155.
- Weitzman, B. 1963. A figural aftereffect produced by a phenomenal dichotomy in a uniform contour. *Journal of Experimental Psychology* 66:195-200.
- Weitzman, B. 1966. Figural aftereffect produced by a phenomenal dichotomy in a uniform contour. *Journal of Experimental Psychology* 71:781-783.
- Willems, E. P. 1967. Nonstimulus and nonretinal mechanisms in figural aftereffects. *Journal of Experimental Psychology* 74:452-454.
- Yoshida, T. 1953. An experimental study of figural aftereffect. *Japanese Journal of Psychology* 23:235-238, 287-288.
- Yoshida, T. 1955. Figural aftereffects as a function of the brightness of figure and ground. Paper read at the 19th Annual Meeting of the Japanese Psychological Association.
- Yoshida, T. 1960. Figural aftereffects as a function of the brightness ratio between the inspection figure and its surrounding field. In *Studies in psychology in commemoration of Professor Matsusaburo Yokoyama's birthday*, Tokyo: Keio University, pp. 99-114.

continued from page 350

- Schmitt, Francis O.; Adelman, George; Melnechuk, Theodore; and Worden, Frederic G., eds. *Neurosciences Research Symposium Summaries. Vol. 5, An Anthology of Work Session Reports from the Neurosciences Research Program Bulletin*. Cambridge, Mass.: The M.I.T. Press, 1972. Pp. vii, 566. \$12.50.
- Schneirla, T. C. *Army Ants: A Study in Social Organization*. San Francisco: W. H. Freeman, 1971. Pp. xx, 349. \$12.00.
- Scott, Edward M. *An Arena for Happiness*. Springfield, Ill.: Charles C. Thomas, 1971. Pp. xiii, 124. \$9.25.
- Segal, Julius, ed. *The Mental Health of the Child: Program Reports of the National Institute of Mental Health*. Bethesda, Md.: Public Health Service, 1971. Pp. ix, 587.
- Segal, Sydney Joelson, ed. *Imagery: Current Cognitive Approaches*. New York: Academic Press, 1971. Pp. xiii, 137. \$7.50.
- Shibles, Warren A. *Metaphor: An Annotated Bibliography and History*. Whitewater, Wis.: The Language Press, 1971. Pp. xiv, 414. \$12.50.
- Silverstein, Michael, ed. *Selected Writings of William Dwight Whitney*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xxxv, 300. \$12.50.
- Siu, R. G. *The Portable Dragon: The Western Man's Guide to the I Ching*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. xv, 463. Paperback, \$1.95.
- Stotland, Ezra; Sherman, Stanley E.; and Shaver, Kelly G. *Empathy and Birth Disorder: Some Experimental Explorations*. Lincoln: University of Nebraska Press, 1971. Pp. 197. \$8.50.
- Stringer, Lorene. *The Sense of Self: A Guide to How We Mature*. Philadelphia: Temple University Press, 1972. Pp. ix, 170. \$8.50.
- Stuart, Richard B., and Davis, Barbara. *Slim Chance in a Fat World: Behavioral Control of Obesity*. Champaign, Ill.: Research Press, 1972. Pp. vii, 240.
- Task Force on Community Mental Health. *Issues in Community Psychology and Preventive Mental Health*. New York: Behavioral Publications, 1971. Pp. xii, 161. \$8.95; paperback, \$4.95.
- Thomas, Alexander, and Sillen, Samuel. *Racism and Psychiatry*. New York: Brunner/Mazel, 1972. Pp. xiii, 176. \$7.50.
- Tobach, Ethel; Aronson, Lester R.; and Shaw, Evelyn, eds. *The Biopsychology of Development*. New York: Academic Press, 1971. Pp. xix, 593. \$22.50.
- Van Der Geer, John P. *Introduction to Multivariate Analysis for the Social Sciences*. San Francisco: W. H. Freeman, 1971. Pp. xi, 293. \$9.50.
- Waters, Maurice. *The United Nations: International Organization and Administration*. New York: Macmillan, 1971. Pp. vii, 583.
- Weekes, Calire. *Peace from Nervous Suffering*. New York: Hawthorn Books, 1972. Pp. x, 206. \$6.95.
- Welford, A. T. *Fundamentals of Skill*. Troy, Mo.: Harper & Row, 1971. Pp. 426. Paperback.

continued on page 392

FACTORS AFFECTING MEMORY SPAN IN BINARY AND OCTAL RESPONDING

Louis Buffardi

Kansas State University

The stimuli were 24 two-alternative elements presented ungrouped, with pauses (temporal grouping), or to different senses (modality grouping), or in a combination grouping. With binary responding, the subjects' recall of the elements was best when there was no stimulus grouping of any kind; with octal responding, when stimulus grouping was by threes. The effects of stimulus-stimulus compatibility, of rate of presentation, and of modality are also considered.

Miller (1956) has reviewed a large body of literature showing that memory span has a fixed capacity of about seven elements—digits, letters, or words. Since the amount of information contained in the individual elements of the test material does not affect memory span, more total information can be retained by increasing the amount of information per element. This regrouping process is known as 'chunking.'

Chunking may be brought about by stimulus grouping or by response grouping. Physically grouping (by time, space, or sense modality) the stimuli in presentation is termed 'stimulus grouping.' Instructing the subject to mentally group the stimuli himself is termed 'response grouping.' This distinction between stimulus and response grouping is one of experimental procedure; no inferences should be drawn with regard to the point in cognition at which the subject actually recodes the stimuli. Furthermore, whether chunking in fact aids performance is a function of at least the method of stimulus or response grouping, the presentation rate of the elements, and the compatibility among these factors.

With respect to the method of *stimulus grouping*, it has been shown that a temporal grouping, by inserting pauses between stimuli, aids memory for sequential presentation of two-alternative elements (Pollock and Johnson, 1965; Parkman and Gregg, 1969). But Klemmer (1964) found that a spatial grouping of simultaneously flashed stimulus lights did not improve performance. And Ryan (1969) compared

temporal grouping (inserting pauses) to nontemporal grouping (inserting brief tones) in the auditory presentation of digit sequences. Temporal grouping of these stimuli significantly reduced errors, but nontemporal grouping was only slightly superior to ungrouped presentation. Thus, different methods of stimulus grouping vary in their influence on performance.

There are two methods of *response grouping*. The informal approach is simply to instruct the subject to rehearse the stimuli by twos, threes, or so on; no transformation of the stimuli is suggested. Parkman and Gregg (1969) found that subjects who had not received any grouping instruction took nearly twice as many trials to learn a sequence of two-alternative elements than did subjects who were given informal instruction. Similarly, Ryan (1969) found fewer errors with informal instruction than with no instruction.

The second response-grouping method is to give instructions in a formal coding system. Here the subject is instructed to transform the stimuli (usually binary digits) into a shortened notation, such as an octal code (Pollack and Johnson, 1965; Klemmer, 1964). In an unpublished study cited by Miller (1956), Smith used an octal-recoding procedure and found that memory span for the digits tripled. Thus, when employed separately, both stimulus and response grouping may aid performance.

The *rate of presentation* is also a factor. In general, rapid presentation impairs performance in memory span and learning experiments. This is especially characteristic of studies involving a formal response-grouping code. A transformation (from binary digits to octal coding, for example) takes time. Unless there is enough time to make this transformation, coding cannot aid performance. Pollack and Johnson found that at a fast presentation rate (2.8 elements/sec) there was little improvement in memory span after octal-coding instructions, although there was a large improvement at the slow presentation rate (.7 elements/sec).

Further, while the effect of *modality* has not been directly examined in recoding studies, there are strong indications that individual modalities would have an effect on performance, particularly at the faster presentation rates. At fast rates, repeating temporal patterns were identified fastest in the auditory modality, intermediate in the haptic, or tactual, modality, and slowest in the visual modality (Garner and Gottwald, 1968; Handel and Buffardi, 1968). In these studies, pattern identification in the visual modality was 50% slower than in the auditory modality.

Finally, Fitts and Seegar (1963) state that the degree of *compatibility* is at a maximum when the reconstruction or regrouping processes are at a minimum. Stimulus-response (S-R) compatibility is maximum when the stimulus elements are grouped into the same number of units as required by the response code. Stimulus-stimulus (S-S) compatibility is maximum when the stimulus elements are identically grouped by each method of stimulus grouping (by an alternation of modality after three elements and a temporal pause after three elements, for example). Pollack and Johnson (1965) concluded that response grouping is most effective when the format of the stimulus grouping is compatible with it. (But, as noted above, only slight improvement was found at fast presentation rates.)

In sum, past research has indicated several factors—stimulus grouping, response grouping, presentation rate, S-R compatibility—that influence performance in a memory-span task. However, no study has systematically investigated the importance of each of these factors, nor have the effects of S-S compatibility or modality on them been studied. Thus, the present study sought to compare these factors and to further investigate their interrelation.

METHOD

—Subjects and basic task— Eight female undergraduates at Kansas State University were paid by the hour for serving in the experiment. Each observed sequences of 24 stimulus events. The stimulus events themselves were two-alternative, or binary, consisting of a right or left stimulus event, and the sequences of them were presented without any grouping at all, grouped by modality only, grouped by temporal breaks in the sequence, or grouped by a modality-temporal combination. After each stimulus sequence, the subject was instructed to write down as much of the sequence as she could remember. She was to use a 0 to represent a right stimulus and a 1 to represent a left stimulus in the binary-response conditions, putting those numbers in 24 boxes arranged in a line across each answer sheet; in the octal-response conditions, she was instructed to use the octal equivalents, from 0 through 7, putting them in the first 8 boxes on the answer sheet (in other words, 'chunking' her responses to the elements by threes).

—Apparatus—The stimulus events could occur in different modalities: vision, audition, or touch (the haptic modality), each of which was represented by two perceptually different stimuli, one to the subject's left, the other to her right. The visual stimuli were two 2.5-cm panel lights (one red, one green) mounted 15 cm apart on a Masonite board positioned 1.2 m in front of the subject. The brightness of each light was .25 ftL. The auditory stimuli were two tones (one 1,200 Hz, one 3,000 Hz). The intensity of both

tones was 65 db sound-pressure level, and they were presented separately to the left and right ears by stereo headphones. The haptic stimuli were two vibrators (one more intense one powered by a 60-Hz, 6.3-v source; the other powered by a 60-Hz, .5-v source), on the top of each of which the subject placed an index finger.

The stimulus patterns were punched on Mylar tape and read out by a tape reader that activated a set of relays to switch the stimuli on and off. Each stimulus element (a right or left) was activated for the first half of the interval between the onset of successive stimuli. The experimenter and equipment were in a small room next to the test room. White noise (50 db) was used to mask any equipment noise.

—Stimulus grouping—The sequences of stimulus events that were presented to each subject were arranged by two methods: by temporal grouping and by modality alternations. *Temporal grouping* consisted of the experimenter's having inserted 'pauses' after a given number of stimulus elements, so that the elements were either presented continuously without any temporal grouping (T_0), grouped by twos (T_2), or grouped by threes (T_3). *Modality grouping* consisted of the experimenter's having alternated the modality in which the elements were presented. This way the elements either were presented entirely in one modality, that is, with no modality grouping (M_0), or were alternated among the three modalities.¹ If the modalities alternated, the modalities could alternate after every element (M_1), after every two elements (M_2), or after every three elements (M_3).

It is possible to use both methods simultaneously. In Table 1, the 12 main stimulus conditions are listed, each condition representing a unique combination of temporal and modality grouping.² As can be seen there, in one condition (the first) neither method grouped the stimuli, in five conditions only one method grouped the stimuli, and in six conditions both methods grouped the stimuli. The last category can be further subdivided into two conditions in which both methods grouped by using the same number of elements (high compatibility) and four conditions in which each method grouped by using a different number of elements (low compatibility). This distinction between levels of S-S compatibility is elaborated, and extended to levels of S-R compatibility, in the discussion section.

All told, there were 216 sequences of stimulus events, each sequence generated from a table of random numbers within the restraints (a) that right and left stimulus elements both occur at least six times in a sequence and (b) that in dividing the 24-element sequence into eight groups of three elements, the same configuration of three stimulus elements (say, three lefts) not occur more than twice in a sequence, and never consecutively (say, six lefts). The 216 sequences so formed were divided into 27 blocks of eight sequences each and the blocks randomly assigned to conditions within the restriction that at least 18 blocks come between repetition of a given block to a given subject.

—Rate of presentation—Two overall rates of stimulus-sequence presentation were used: *slow* (1 stimulus element/sec) and *fast* (3.125 stimulus elements/sec). Total time was held constant, so that a 24-element sequence took 24 sec at the slow rate and 7.68 sec at the fast rate for all 12 conditions.

Since temporal grouping required adding pauses, the 12 conditions were equated for overall presentation rates (fast and slow) in the following manner.

For nontemporally grouped conditions (T_0), the stimuli were presented at 1 element/sec for the slow rate and 1 element/.32 sec for the fast rate. For grouping by twos (T_2) at the slow rate, each two-element group took 1 sec followed by a 1-sec pause; at the fast rate each two-element group took .32 sec followed by a .32-sec pause. For grouping by threes (T_3) at the slow rate, each three-element group took 1.5 sec followed by a 1.5-sec pause; at the fast rate each group took .48 sec followed by a .48-sec pause. Consequently, for all temporally grouped conditions, half of the total presentation time was used for presenting the stimuli, while the remaining half was used for pauses that separated the elements into groups.³

—Response grouping—The eight subjects were tested for approximately 75 min a day for 15 days in the binary-response condition (see the section on subjects and task, above); then they were taught the octal code for 3 days and retested for 15 days in the octal-response condition. The latter can be more fully described.

During the first day of octal-coding instruction, the code was described and each subject was given a list of the possible combinations of three stimulus elements (each a right or a left) and their octal equivalents. She studied this list for 5 min and was then tested by means of flash cards. After she gave the correct octal response 48 consecutive times, the session was terminated. On the second day, she was tested for 60 min on the experimental apparatus. Here, groups of three stimulus elements were presented in the individual modalities at both rates, the subject instructed to verbally report the appropriate octal number.

On the third, and final, day of coding instruction, each subject practiced for brief periods on each of the conditions that involved modality alternations for a total of 60 min. Here, as before, she was instructed to respond in octal terms after each group of three elements—except when the elements were grouped in pairs, when she was to observe three pairs of elements and then respond with two octal numbers. After three days of instruction, the subjects were well versed in the octal code (making very few mistakes after the first few trials on the first day of training) and had received equivalent amounts of familiarization with all experimental conditions.

—Scoring—The average number of errors per 24-element sequence was the measure of performance. Since the subjects were told to leave blanks for elements they could not recall, each blank space was counted as .5 error for the binary-response condition. For the octal-response condition, an error in octal notation was treated as equivalent to a response of three binary elements. Reconversion to binary elements yielded 1.71 binary errors per error in octal notation, on the assumption that all errors in octal were equally likely. This is probably a conservative conversion (see Pollack and Johnson, 1965). Each blank in the octal condition was counted as equivalent to 1.5 binary errors.⁴

—Statistical procedures—A problem arose in determining the effect of octal

coding on subsequent performance. Since the octal responding always followed the binary responding, any improvement in the performance of the former might be due to a general learning effect, not to the octal responding itself. For example, in Figure 1 (overall performance curve) the mean improvement for the octal-response condition might be due to a learning effect. However, a simple t test between mean performance during the binary and octal (response) conditions would be significant. Hence, the appropriate comparison is not between binary and octal coding but between (a) an estimate of the octal-condition performance if octal coding had no effect and (b) the actual octal-condition performance.

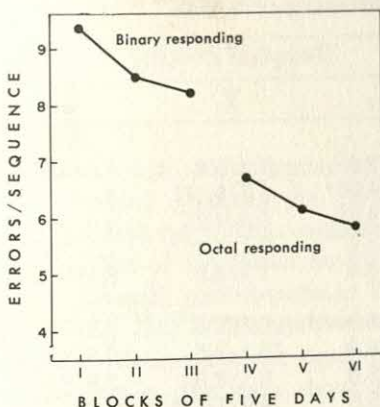


Fig. 1. Average number of errors made per 24-element sequence, averaged across all regular conditions and presentation rates, for each block of five days

Box and Tiao's (1965) Bayesian model for an interrupted, nonstationary time series provides this comparison. The model, by means of a modified t test, answers the question of whether the observations after the instruction in octal coding were simply a continuation of the time series of the binary-condition observations or whether they shifted up or down from that general level. The model is a special case of the integrated moving-average process, a process which has been used previously with success in representing a nonstationary time series. The model regards the system as being subjected to periodic shocks (in this case, the learning that takes place over time), a proportion of each shock remaining in the series through time. Thus it is possible to estimate future observations (to extrapolate them from the binary-condition data) by an exponentially weighted average of all previous observations. Hence, the technique assesses the statistical stability of an alteration in the course of a time series. Campbell (1969) has discussed the general suitability of the Box and Tiao model for time-series data.

RESULTS

For both binary and octal (response) conditions, the number of errors per sequence was averaged across subjects and presentations. The mean number of errors per sequence for each condition at each presentation rate is shown in Table 2. All findings discussed below are based on *F* ratios and were significant at least at the .05 level, except where noted.

Table 2. Number of errors made per 24-element sequence, averaged across subjects and presentations, for the binary/octal response conditions at slow and fast presentation rates

Modality grouping	Temporal grouping			Mean
	T ₀	T ₂	T ₃	
1 element/sec				
M ₀	7.3/5.2	7.3/5.8	7.8/3.3**	7.5/4.8
M ₁	8.6/4.0**	8.5/5.0**	8.3/3.3**	8.5/4.1
M ₂	7.9/4.8*	7.6/4.2*	8.4/4.0**	8.0/4.3
M ₃	8.1/3.4**	8.1/4.4**	7.9/3.1**	8.0/3.6
Mean	8.0/4.4	7.9/4.9	8.1/3.4	8.0/4.2
3.125 elements/sec				
M ₀	8.6/8.3	9.1/9.4	9.2/7.1	9.0/8.3
M ₁	10.0/8.9	10.3/9.5	10.2/9.1	10.2/9.2
M ₂	9.2/8.6	9.1/8.9	9.8/8.6	9.4/8.7
M ₃	8.9/5.8**	9.9/7.7**	9.0/6.3**	9.3/6.6
Mean	9.2/7.9	9.6/8.9	9.6/7.8	9.5/8.2

*Significant improvement due to octal coding at .05 level.

**Significant improvement due to octal coding at .01 level.

Binary responding

With binary responding, stimulus grouping produced poorer performance. The M₀T₀ condition, in which there was no grouping, produced the best performance at both presentation rates.

With respect to *individual modality*, it will be recalled that three conditions (M₀T₀, M₀T₂, and M₀T₃) were presented in auditory, haptic, and visual modalities. The interaction of individual modality and condition was not significant, so that it is valid to sum across conditions for modality effects and vice versa. A significant main effect was found for modalities, with performance in the auditory modality best (7.9), that in the visual modality intermediate (8.1), and that in the haptic modality worst (8.5).

The results on *rate of presentation* show that performance at the slow rate was significantly better than at the fast rate (8.0 to 9.5 errors). Presentation rate did not interact with methods of stimulus grouping; the difference between slow and fast rates was constant across conditions.

The data on *modality grouping* show a significant difference between those conditions in which patterns were presented in one modality and those in which the modality alternated. No modality grouping (M_0) was best (8.2), M_2 and M_3 were intermediate (8.7), and M_1 produced the poorest performance (9.4). The data on *temporal grouping* show that ungrouped presentation (T_0) produced better performance (8.6) than T_2 (8.7) or T_3 (8.8). These differences between temporal-grouping conditions were not significant.

Octal responding

The best performance with octal responding was produced by those conditions using stimulus grouping by three elements. Condition M_3T_3 was best at the slow rate, while condition M_3T_0 was best at the fast rate. The interaction of *individual modality* and condition again was not significant. Overall, performance in the visual modality was best (6.2), followed by that in the auditory (6.5) and haptic (6.8) modalities. These differences were not significant.

On *rate of presentation*, the data show significantly better performance at the slow rate than at the fast rate (4.2 to 8.2 errors). Presentation rate did not interact with methods of stimulus grouping; however, the differences between slow and fast rates were greater for octal responding (4.0 errors) than they were for binary responding (1.5 errors).

For *modality grouping*, modality alternations every three elements (M_3) produced the best performance (5.1), followed by M_0 (6.5), M_2 (6.5), and M_1 (6.6). The last three conditions did not differ among themselves. For *temporal grouping*, T_3 was superior (5.6), with T_0 intermediate (6.1) and T_2 (6.9) worst. Only the difference between T_3 and T_2 was significant.

The overall effect of octal coding itself on recall of the sequences was determined by the Box and Tiao test. In Table 2, improved performance due to octal coding is shown by asterisks. At the slow rate, octal coding was beneficial, especially for conditions in which at least one stimulus grouping was by three elements. For only two conditions (M_0T_0 , M_0T_2) did octal coding not improve performance. However,

at the fast rate it was only when modality grouping was by threes (M_3T_0 , M_3T_2 , M_3T_3) that octal coding significantly improved performance. No significant improvement occurred for any other condition.

DISCUSSION

Individual modalities

Memory span depended on the modality in which the stimuli were presented. Presentation in the auditory (7.2) and visual (7.2) modalities produced fewer errors than presentation in the haptic modality (7.6). However, the effect of modalities was small and, more important, did not interact with stimulus or response grouping. There was a minor interaction with presentation rate: the visual modality (5.9) was easier than the auditory one (6.2) at the slow rate, but the auditory one (8.3) was easier than the visual (8.5) at the fast rate. Haptic presentation was poorest at both rates. This result can be contrasted to the large differences between auditory and visual modalities in the identification of temporal patterns (Garner and Gottwald, 1968; Handel and Buffardi, 1968). The recall of two-alternative stimulus elements is more a cognitive task (of learning and memorizing) than it is one of pattern perception, and this probably minimizes any differences between modalities.

Methods of stimulus grouping

It must be concluded that stimulus grouping can aid short-term memory span only when the method of stimulus grouping is compatible with the method of response grouping (S-R compatibility). This conclusion is most apparent for octal responding, where the three easiest conditions, averaged across presentation rate, were M_0T_3 , M_3T_0 , and M_3T_3 ; furthermore, M_3 and T_3 produced better performance than any other grouping.

For binary responding, the ungrouped condition (M_0T_0) produced the best performance. Previous work (Pollack and Johnson, 1965) reported that when subjects are left on their own, they form their own response codes. Similarly, in the present experiment, the subjects said that they organized the sequence by runs of identical elements (say, 'four lefts, three rights, one left'). Thus, the condition in which

there was no stimulus grouping was the condition most compatible with the binary-response code—either modality or temporal grouping would tend to break up these runs and interfere with this method of organization (Handel, 1969).

That stimulus grouping in the binary-response conditions hindered performance is at odds with previous research (Ryan, 1969). However, the stimuli in previous work were strings of random digits. Temporal pauses did not break up structure in the string, but created structure by grouping. This improved performance. In contrast, in the present experiment the pauses tended to break up the structure of runs in the strings.

The effects of temporal and modality grouping differed for binary and octal coding. With binary responding, temporal grouping of stimuli without modality grouping (M_0T_2 , M_0T_3) produced fewer errors than modality grouping without temporal grouping (M_2T_0 , M_3T_0), although the differences were not great (8.4 to 8.5 errors). But with octal responding, the ranking was reversed: modality grouping produced fewer errors than temporal grouping (5.7 to 6.4 errors). Possibly a transformation code affects the relative efficiency of stimulus-grouping methods. The inferiority of temporal grouping with octal responses may also have been due to the faster rate of presentation of the stimuli (faster because of the pauses added within the same amount of total time). But if increased presentation rate is responsible, it is not clear why temporal grouping was superior to modality grouping with binary responses.

The control conditions for temporal grouping, conditions which allowed total time to increase (see note 3), yielded better performance than the modality-grouping conditions. This again suggests that the faster presentation rates were responsible for temporal-grouping inferiority. But the superior performance in the temporal control conditions might also have been due to the increased amount of time to process the sequences. Because of the unavoidable confounding of presentation rate and total time, it is impossible to resolve the question of whether temporal or modality grouping was the more efficient.

The difference in recall between slow and fast presentation rates was nearly three times as great for octal than for binary responses (4.2 versus 1.5 errors). Octal transformations take time, and the subjects may not have been able to keep up at fast presentation rates. This difference was minimized for modality grouping by threes, a condition which would tend to make recoding easy.

S-S and S-R compatibility

Each stimulus grouping can be classified as to the degree of S-S compatibility. High compatibility was when both modality alternation and temporal pauses grouped the same number of stimuli (M_0T_0 , M_2T_2 , M_3T_3), medium compatibility was when only one of these methods grouped the stimuli (M_0T_3 , M_3T_0 , M_0T_2 , M_2T_0 , M_1T_0), and low compatibility was when each method grouped into a different number of elements (M_1T_2 , M_1T_3 , M_2T_3 , M_3T_2).

For octal responding each condition can similarly be classified as to S-R compatibility. Since the octal code is based on three elements, high S-R compatibility occurred when at least one method grouped the stimuli by threes (M_0T_3 , M_1T_3 , M_2T_3 , M_3T_0 , M_3T_2 , M_3T_3) and low S-R compatibility was when neither method grouped by threes (M_0T_0 , M_0T_2 , M_1T_0 , M_1T_2 , M_2T_0 , M_2T_2).

For octal responding, averaged across presentation rates, an analysis of variance found that S-R compatibility accounted for 75% of the variance (1 of 11 *df*) and S-S compatibility accounted for 13% of the variance (2 of 11 *df*). The means are shown in Figure 2. The same analysis was done for binary responding. Here, although S-R compatibility is meaningless, S-S compatibility accounted for 61% of the variance and S-R compatibility accounted for 0%.

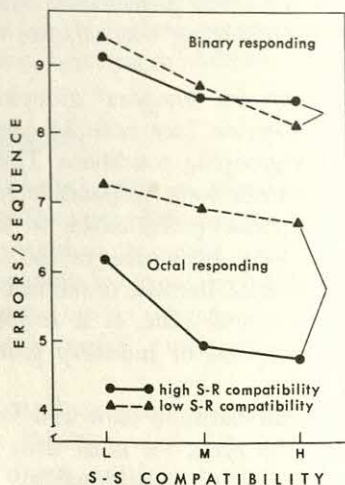


Fig. 2. Average number of errors made per 24-element sequence, averaged across presentations and rate, for the conditions comprising the two levels of S-R compatibility at each level of S-S compatibility

To determine the effect of S-S and S-R compatibility on the facilitory effect of octal coding, the Box and Tiao technique was applied. For low S-R compatibility, no level of S-S compatibility showed significant improvement in performance. However, for high S-R compatibility, significant improvement was found for all levels of S-S compatibility. Therefore, the present study replicated the results of Pollack and Johnson (1965) by finding S-R compatibility crucial to performance when an octal response is used. Furthermore, it showed that S-S compatibility had a significant but relatively minor role, taking on importance only within levels of S-R compatibility. In addition, octal coding facilitated performance chiefly at the slow rate.

Selective attention and filtering

There was strong evidence that when both stimulus-grouping methods are used simultaneously, subjects selectively filter out the grouping method that reduces performance and attend to the grouping method that gives the best performance. With binary responses all stimulus grouping was detrimental, but when possible, the subjects filtered out the stimulus groupings that most reduced recall. For instance, both temporal grouping and modality grouping increased errors here, but these effects were not additive; redundant grouping (M_2T_2 , M_3T_3) was equal to temporal grouping, its best component. Also, recall for those binary-response conditions in which neither modality nor temporal grouping occurred was better than recall for those conditions in which both modality and temporal grouping occurred (7.8 to 8.2 errors at the slow rate, Mann-Whitney U , $p < .15$; 9.2 to 9.7 at the fast rate, Mann-Whitney U , $p < .07$).

The conclusion that subjects selectively attend to the best type of stimulus grouping was supported by the octal-response results. Here grouping aided performance; modality grouping was superior to temporal grouping, and redundant grouping was at least as good as modality grouping. Furthermore, consistently better performance was found with high S-R compatibility (when either modality or temporal grouping was by threes), even though all but one of these conditions also had a component that did not group by threes (3.6 to 4.8 errors at the slow rate, Mann-Whitney U , $p < .02$; 7.4 to 8.9 errors at the fast rate, Mann-Whitney U , $p < .05$). The subjects attended to the method of stimulus grouping that was compatible with the octal code and thereby improved their recall.

Notes

This study is based on a dissertation submitted to Kansas State University in partial fulfillment of the requirements for the Ph.D. The author thanks Dr. Stephen Handel for his helpful guidance. This research was supported in part by NIH Grant MH 15969 and was carried out during the author's tenure as an NIMH trainee (5t01 MH08359). Revision received for publication July 1, 1971.

1. When all the stimulus elements were presented in one modality, each modality was used an equal number of times; when the modalities alternated, the six possible orders of modality alternation were counter-balanced.

2. For design purposes only, the 12 conditions were divided into six sections: (a) M_0T_0 , with auditory, haptic, and visual subconditions; (b) M_0T_2 , with auditory, haptic, and visual subconditions; (c) M_0T_3 , with auditory, haptic, and visual subconditions; (d) M_1T_0 , M_2T_0 , and M_3T_0 ; (e) M_2T_2 , M_3T_3 , and M_1T_3 ; and (f) M_2T_3 , M_3T_2 , and M_1T_2 . Every day each subject was presented with four of these six sections, so that every three days all conditions appeared twice for each subject. For each condition in a section, four sequences of 24 elements were presented at each rate. The order of presentation of the six possible sections was also balanced for each subject with the restriction that the other five sections appear at least once before any single section repeated. Within each section the order of conditions was counterbalanced.

3. To control for this manipulation, two extra stimulus conditions paralleling M_0T_2 and M_0T_3 were added to those shown in Table 1. These conditions used the same presentation rate but let total time per sequence vary. Each pause was one element long (1 sec at the slow rate, .32 sec at the fast rate). Thus, the total time for a 24-element sequence at a slow rate for control M_0T_2 was 36 sec; for control M_0T_3 , it was 32 sec. These stimulus conditions were presented three times each for both binary and octal responding for each subject, on days 1-2, 7-8, and 14-15. To keep the length of the experimental day to approximately 75 min when one of these control conditions occurred, it substituted for one of the regular sections (see note 2, above). Hence, each regular stimulus condition appeared nine times for both binary- and octal-response conditions, with the conditions balanced across each block of five days.

4. Other conversions are possible. For example, errors and blanks could both count as 1.71 errors or both as 1.5 errors. These conversions were calculated, but they did not affect the pattern of results.

References

- Box, G. E. P., and Tiao, G. C. 1965. A change in level of a non-stationary time series. *Biometrika* 52:181-192.
- Campbell, D. T. 1969. Reforms as experiments. *American Psychologist* 24:409-429.
- Fitts, P. M., and Posner, M. I. 1967. *Human performance*. Belmont, Calif.: Brooks/Cole.

- Fitts, P. M., and Seegar, C. M. 1953. S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology* 46:199-210.
- Garner, W. R., and Gottwald, R. L. 1968. The perception and learning of temporal patterns. *Quarterly Journal of Experimental Psychology* 20:97-109.
- Handel, S. 1969. Perception and learning of temporal patterns. Paper presented at the meeting of the American Association for the Advancement of Science, Boston, December.
- Handel, S., and Buffardi, L. 1968. Pattern perception: Integrating information presented in two modalities. *Science* 162:1026-1028.
- Klemmer, E. T. 1964. Does recoding from binary to octal improve the perception of binary patterns? *Journal of Experimental Psychology* 67:19-21.
- Miller, G. A. 1956. The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63:81-97.
- Neisser, U. 1966. *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Parkman, J., and Gregg, L. W. 1969. Effects of temporal grouping and instruction on serial acquisition of binary numbers. Paper presented at the meetings of the Midwestern Psychological Association, Chicago, May.
- Pollack, I., and Johnson, L. B. 1965. Memory-span with efficient coding procedures. *American Journal of Psychology* 78:609-614.
- Posner, M. I. 1963. Immediate memory in sequential tasks. *Psychological Bulletin* 60:333-349.
- Royer, F. L., and Garner, W. R. 1966. Response uncertainty and perceptual difficulty of auditory patterns. *Perception and Psychophysics* 1:41-47.
- Ryan, J. 1969. Grouping and short-term memory: Different means and patterns of grouping. *Quarterly Journal of Experimental Psychology* 21:137-147.

continued from page 376

- Weinberg, Gerald M. *The Psychology of Computer Programming*. New York: Van Nostrand Reinhold, 1971. Pp. xv, 288. \$9.50.
- Wesley, Frank. *Childrearing Psychology*. New York: Behavioral Publications, 1971. Pp. viii, 243. \$8.95; paperback, \$4.95.
- Wheelis, Allen. *The End of the Modern Age*. New York: Basic Books, 1971. Pp. 129. \$5.95.
- White, Owen R. *A Glossary of Behavioral Terminology*. Champaign, Ill.: Research Press, 1971. Pp. xxv, 220.
- Wiener, Daniel N. *Classroom Management and Discipline*. Itasca, Ill.: F. E. Peacock, 1972. Pp. 138. \$5.00; paperback, \$3.50.
- Wike, Edward L. *Data Analysis: A Statistical Primer for Psychology Students*. Chicago: Aldine-Atherton, 1971. Pp. x, 256. \$8.50; paperback, \$5.95.
- Wilson, Colin. *The Occult: A History*. New York: Random House, 1971. Pp. 601. \$10.00.
- Wilson, Edward O. *The Insect Societies*. Cambridge, Mass.: Harvard University Press, 1971. Pp. 548. \$20.00.
- Windle, William F. *Physiology of the Fetus: Relation to Brain Damage in the Perinatal Period*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xi, 160. \$10.50.
- Wortis, Joseph, ed. *Mental Retardation*. Vol. 3. New York: Grune & Stratton, 1971. Pp. ix, 243. \$15.75.
- Wright, George N., and Reagles, Kenneth W. *The Economic Impact of an Expanded Program of Vocational Rehabilitation*. Madison: University of Wisconsin Regional Rehabilitation Research Institute, 1971. Pp. xv, 69.
- Zusman, Jack, and Davidson, David L., eds. *Organizing the Community to Prevent Suicide*. Springfield, Ill.: Charles C Thomas, 1971. Pp. ix, 97. \$6.25.

COLOR RECOGNITION AND WORD FREQUENCY

Ruth F. Deich

California Department of Mental Hygiene, Pacific State Hospital

College students learned a list of six paired associates with high- or low-frequency color names as stimuli and two-digit numbers as responses, then were shown color cards matching the color names and asked to recall the numbers. With rehearsal and color blindness controlled, the subjects learning high-frequency names had significantly better recall, confirming that color memory is a function of linguistic codability, not perceptual experience.

It has been found that the ease of recognizing or remembering a color is directly related to frequency of use of that color name. If Brown and Lenneberg (1954) and Lenneberg (1961) are correct, this phenomenon is a linguistic one. If Horowitz, Norman, and Day (1966) are correct, it is perceptual in nature, since "the most codable color names (high-frequency words) are probably used to label the most frequent sensory experiences" (p. 13). If the phenomenon is in fact based on sensation, then using a high-frequency word like 'blue' should result in no better recognition than a low-frequency synonym like 'azure,' since both have approximately the same frequency of sensation. If it is based on frequency of codability, the word blue should result in better recognition than the word azure. It was hypothesized that color recall is a function of codability and not of frequency of sensory experience.

METHOD

Two groups of college students learned a list of six paired associates by the method of anticipation to the criterion of one correct reproduction. Before list presentation, the subjects were told that this was an experiment to compare learning rates among different groups.

A Kodak 800 projector with a tachistoscopic attachment permitted control of exposure time. The latter was 1 sec for each slide and 1 sec for the blank interval between slides. Exposure was the same regardless of whether the slide showed only the stimulus or the complete stimulus-response pair.

Stimuli were color names. Responses were two-digit numbers. The control group ($N = 21$) was given well-known, high-frequency color names

in the following stimulus-response pairs: RED-29, BLUE-51, GREEN-73, BLACK-16, PINK-87, ORANGE-34. The experimental group ($N = 21$) was given as stimuli four low-frequency dictionary synonyms: CRIMSON, AZURE, JADE, and EBONY. The other two color names, PINK and ORANGE, remained the same, as did all the response numbers. Pair orders were counterbalanced in both groups.

Immediately after learning, the subjects were given the Dvorine pseudo-isochromatic plates (1953) to control for color-vision abnormalities and to prevent rehearsal. Then they were shown six color cards.¹ Each of these cards (which were the same for all of the subjects) had a colored tissue paper that matched one of the previously learned color names, and the cards were presented in a different order than that of the color names on the learning list. The color cards were exposed for 2 sec each, and the subject was asked to recall the appropriate associated response number.

The cards were $\frac{1}{8}$ in. thick, with blanks between the color cards. This thickness made the cards easy to hold and to drop. The 2-sec intra- and interitem exposures were timed with a stopwatch and metronome. A constant exposure rate was attained, as the experimenter (a research assistant with no knowledge of the experimental hypothesis) was well practiced on a pilot group.

After recall, each subject was given a sheet of paper showing the learning-task color names and nine differently colored dots (including yellow and brown and gray). Each subject was required to match each color name with the colored dot most closely approximating the name. This procedure controlled for the subject's not knowing the meaning of any particular color name.

RESULTS

The data from 36 of the original 42 subjects were available for analysis. One in each group was eliminated on the basis of red/green color blindness, and four additional subjects in the experimental group were eliminated because they did not know one or more of the synonymous, infrequent color names.

Table 1. Mean ($\bar{X}$) trials to learn and two-digit response numbers recalled

	Trials to learn			Numbers recalled	
	Two identicals	Four synonyms	Total list	Four synonyms	Total list
Experimental group ($N = 16$)	2.5	3.4	3.4	2.1°	3.6°
Control group ($N = 20$)	2.6	3.4	3.5	3.6°	5.3°

° $p < .01$.

Both groups required the same number of trials to learn regardless of whether one considers the total list of six pairs ($X=3.4, 3.5$), the four synonymous pairs ($X=3.4, 3.4$), or the two identical pairs ($X=2.5, 2.6$); see Table 1.² The control group recalled significantly more color names than the experimental group, with t of 4.84 for the total list and t of 4.91 for the synonymous pairs ($df = 34, p < .01$).³

DISCUSSION

The results clearly show that color memory is a function of linguistic codability and not of perceptual experience. The equal learning by both groups and the significantly greater retention for high-frequency color names (by the control group) indicate that Brown and Lenneberg (1954) were correct in calling the phenomenon a linguistic one. Thus, a more frequently used color name, such as 'red,' is more readily available than a less frequently used one, such as 'crimson.'

It may be objected that the colors on the recall swatches were not actually crimson, azure, and so on for the experimental group. However, both the colors on the recall swatches and the colors controlling for knowledge of color names had been labeled by the manufacturer with the infrequent color names. Idiosyncratic differences can be ruled out, since all subjects were able to match infrequent color names in the colored swatches shown (see results, above).

The process of color memory in the present experiment may be conceptualized as follows. On the learning task, a printed symbol designating a color cognitively represents a percept of some wavelength range. On the retention task, the subject sees a color of a specific wavelength. He must recall which of the six previously seen color-name stimuli (representing clearly distinct colors) most closely represents the actual color shown: is it red/crimson, or blue/azure, or what? And once he knows the color-name stimulus, he then has to recall the associated response.

Thus, essentially two associations were involved in the retention task: color with color name, and color name with response number. Since the association of name with number was learned at the same rate for experimental and control groups, the differences in retention may be sought in the first link of the chain, color $\rightarrow$ color name $\rightarrow$ response number. If prior frequency of sensory experience were the explanatory factor, there should have been no differences in retention between the two groups, since the same hue was shown to both. But

the difference in favor of the control group indicates that color memory is a function of codability per se, and not codability as dependent on frequency of sensory experience.

Notes

The opinions or conclusions stated in this paper are those of the author and are not to be construed as official or as necessarily reflecting the policy of the Department of Mental Hygiene. Revision received for publication July 7, 1971

1. A projector was not used because there was difficulty obtaining color fidelity on the slides.

2. The subjects had only to learn the response members, which were the same for both groups. This technique avoided possibly greater difficulty with the more unfamiliar color names. The same responses for both groups, together with the brevity of the list, presumably accounted for the equal learning.

3. The recall measure was a two-digit number. A difference in number recall by definition reflects a difference in color-name recall.

References

- Brown, R. W., and Lenneberg, E. H. 1954. A study in language and cognition. *Journal of Abnormal and Social Psychology* 49:454-462.
- Dvorine, I. 1953. *Dvorine pseudo-isochromatic plates*. Baltimore: Waverly Press.
- Horowitz, L. M., Norman, S. A., and Day, R. S. 1966. Availability and associative symmetry. *Psychological Review* 73:1-15.
- Lenneberg, E. H. 1961. Color-naming, color recognition, color discrimination: A reappraisal. *Perceptual and Motor Skills* 12:375-382.

PROBABILITY LEARNING WHEN STIMULI VARY ON MULTIPLE DIMENSIONS

E. Scott Geller and Charles P. Whitman

Virginia Polytechnic Institute and State University

With both location and shape of stimuli varied simultaneously on independent binary probability schedules, the prediction distributions of subjects who had a 70%-30% schedule on each dimension eventually 'matched' the theoretical probabilities, but those of subjects who had a 50%-50% on one dimension underapproximated the theoretical probability for the 70% alternative of the other dimension. Even with the color of the stimuli also varied, subjects' prediction distributions showed they were discriminating all three schedules.

The standard probability-learning experiment first introduced by Humphreys (1939) is a two-choice decision task in which stimulus events are sequentially presented on a predetermined probability schedule and the subjects make a stimulus prediction before each presentation. As the experimental session progresses, the subjects frequently demonstrate 'matching' behavior; that is, the frequency distribution of their stimulus predictions comes to approximate the event probabilities of the presentation schedule (Estes, 1964). The effect is also found when the design is modified to represent a more complex probability situation—when the number of stimulus alternatives is increased beyond two (see the review by Vlek, 1970), when two independent probability schedules are alternated on successive trials (Schipper, 1970), and when the probability of a stimulus alternative is varied intrasequentially in a two-choice and a three-choice decision task (Geller, 1970).

Unlike the previous studies, the present experiments *simultaneously* differentiated the stimulus events along more than one (binary) dimension. Before each stimulus presentation in Experiment I the subjects verbally predicted which of two shapes they expected and in which of two locations it would occur. By independently varying the relative probability of two shapes and two locations, the sensitivity of the subjects' predictions to two simultaneously occurring probability distributions was determined. In Experiment II each stimulus event

represented one of two shapes, two locations, and two colors, so that the ability of subjects to discriminate between three simultaneously occurring probability distributions was studied.

EXPERIMENT I

Method

—Subjects—Volunteers from the introductory psychology classes of Virginia Polytechnic Institute and State University were assigned to three experimental groups of ten males and ten females each.

—Groups—Each group received a different schedule of 200 stimulus presentations in which the probability of an up or down (D) *shape* and a left or right (R) *location* varied independently. Group 70D/70R was presented a 70%-30% distribution of both the shape and the location dimensions in each block of 50 trials (35 down, 15 up shapes; 35 right, 15 left locations). Group 70D/50R was presented a 70%-30% distribution of the shape alternatives (35 down, 15 up shapes per 50-trial block) and a 50%-50% distribution of the location alternatives (25 right, 25 left locations per 50-trial block). Group 50D/70R was presented a 50%-50% distribution of shape (25 down, 25 up shapes per 50-trial block) and a 70%-30% distribution of location (35 right, 15 left locations per 50-trial block). Because the analysis involved only between-group comparisons, counterbalancing of the probabilities of the stimulus alternatives within each dimension was not part of the design.

—Procedure—The subjects performed in individual one-hour sessions. At the start of each session, the prediction task was explained. The subjects were told that the stimulus list was predetermined on punched paper tape and were encouraged to try to make as many correct shape and location predictions as possible. They were given no information about the nature of the stimulus sequence, that is, were not told to expect randomness or patterning. After 10 practice trials, 200 stimulus events (experimental trials) were successively presented; before each presentation the subjects predicted which of the two shapes and two locations would occur next.

Results and discussion

For analysis, the proportion of *down* and *right* predictions for each group was determined for each successive block of 10 trials. As Figure 1 indicates, the subjects' prediction strategies did show a discrimination of two simultaneously presented probability schedules. After 90 trials the prediction distributions of group 70D/70R reached an asymptote that approximated the theoretical probabilities for both downs and rights. The other two groups, both of which received a 70%-30% distribution on one dimension and an equiprobable (50%-50%) dis-

tribution on the other, did not 'match' on the 70%-30% schedule, but after 90 trials their prediction proportions for the 70%-30% dimension were indeed higher for the more probable alternative; for the dimension with equiprobable alternatives their prediction proportions for each alternative fluctuated around .50.

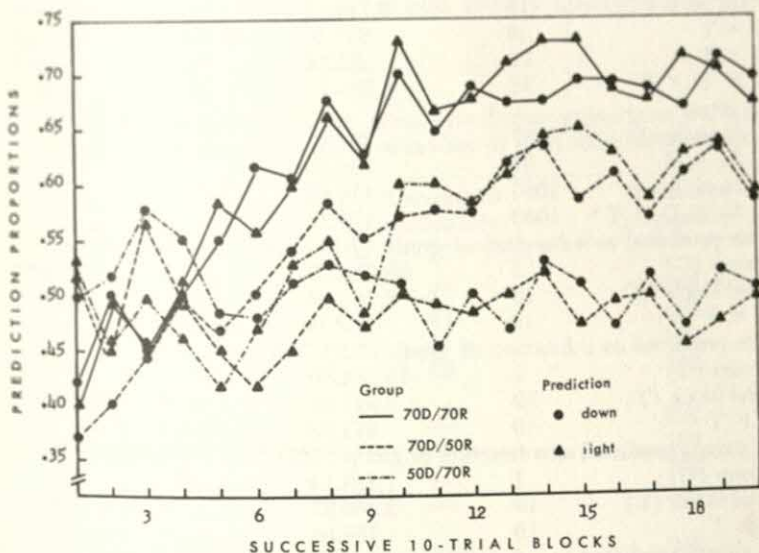


Fig. 1. The proportion of down (D) and right (R) predictions for each schedule group as a function of successive 10-trial blocks

For every trial block after the first 40 trials, the prediction proportion of a 70% alternative (either a down or a right) was higher when the probability schedule for the other dimension was similar (group 70D/70R) than when the probability schedule for the other dimension was dissimilar (groups 70D/50R, 50D/70R). As indicated in Table 1, the analyses of variance comparing the prediction proportions for the 70% alternatives indicated significant group effects between groups 70D/70R and 70D/50R and between groups 70D/70R and 50D/70R. A significant effect of trial block was obtained in each case, and no significant interaction of groups and trial blocks occurred.

Thus, when they received one dimension on a 50%-50% probability schedule, the subjects' tendency to predict the more probable stimulus on the other (70%-30%) dimension was suppressed. Ap-

Table 1. Analyses of variance, Experiment I

Source	df	MS	Error	F
3 (groups) $\times$ 2 (dimensions) $\times$ 20 (trial blocks)				
Group (G)	2	17,993.60	a	21.07**
Dimensions (D)	1	459.38	b	0.58
G $\times$ D	2	7,042.63	b	8.84**
Trial block (T)	19	2,743.81	c	14.10**
G $\times$ T	38	637.10	c	3.27**
D $\times$ T	19	382.54	d	2.15**
G $\times$ D $\times$ T	38	1,264.11	d	7.09**
Error terms				
a. Ss within G	57	853.81		
b. Ss $\times$ D	57	796.30		
c. Ss $\times$ T	1083	194.65		
d. Ss $\times$ D $\times$ T	1083	178.37		
Downs predicted as a function of group (70D/70R versus 70D/50R)				
Group (G)	1	28,084.50	a	32.89**
Trial block (T)	19	2,588.26	c	13.30**
G $\times$ T	19	279.34	c	1.44
Rights predicted as a function of group (70D/70R versus 50D/70R)				
Group (G)	1	13,284.50	a	15.56**
Trial block (T)	19	2,947.00	c	15.14**
G $\times$ T	19	271.08	c	1.39
70% events predicted as a function of group (70D/50R versus 50D/70R)				
Group (G)	1	1,128.13	a	1.32
Trial block (T)	19	2,259.97	c	11.61**
G $\times$ T	19	158.65	c	.82
50% events predicted as a function of group (70D/50R versus 50D/70R)				
Group (G)	1	2.00	a	.002
Trial block (T)	19	1,089.71	c	5.60**
G $\times$ T	19	202.27	c	1.04

** $p < .01$.

parently, subjects simultaneously process two independent probability distributions more appropriately when the two distributions are similar rather than dissimilar. Perhaps interference occurs when the distributions of two independently occurring stimulus dimensions are different, whereas facilitation (of probability information on one stimulus dimension) occurs when the other stimulus dimension is simultaneously varying on an identical distribution.

The effect of a similar distribution on the 'other' stimulus dimension is also evidenced by the comparison in Table 2 between the groups' prediction proportions for their last 100 trials and the theoretical joint probabilities of each stimulus combination. Probability learning is

indicated by the degree of correspondence between each prediction proportion and the appropriate theoretical probability. The interfering effect of a second, dissimilar probability distribution is demonstrated by the underapproximations of the joint probabilities containing a 70% and a 50% stimulus alternative. On the other hand, for the group receiving similar distributions the proportion for the stimulus combination containing both 70% stimuli closely approximated the theoretical probabilities.

Table 2. Comparison between the groups' prediction proportions for the last 100 trials and the theoretical joint probability of each stimulus combination

Stimulus combination	Theoretical probability	Prediction proportion
Group 70D/70R		
down/right	.49	.492
down/left	.21	.189
up/right	.21	.192
up/left	.09	.130
Group 70D/50R		
down/right	.35	.292
down/left	.35	.306
up/right	.15	.192
up/left	.15	.210
Group 50D/70R		
down/right	.35	.306
down/left	.15	.185
up/right	.35	.313
up/left	.15	.199

To determine whether probability information was more readily learned for the shape or location dimension, the prediction distributions for each dimension were compared between groups 70D/50R and 50D/70R by separate analyses of variance. The analysis of variance comparing the prediction distributions for the 70% alternative (down for group 70D/50R and right for group 50D/70R) indicated only a significant effect of trial block; neither the group effect nor the interaction of groups and trial blocks was significant (see Table 1). Likewise, a between-group comparison of the prediction proportions for the 50% alternative (right for group 70D/50R and down for group 50D/70R) demonstrated a significant effect of trial blocks, but again neither the main effect of groups nor the interaction of

groups and trial blocks was significant.¹

This last finding supports the utility of a two-dimensional design. That is, since the data indicate that subjects process shape and location probabilities similarly, the generalization (biasing or interfering) effect of manipulating one probability distribution can be studied with the two-dimensional design. Whereas formerly the effect of an independent variable on probability learning has been studied between groups or stimulus sequences, the two-dimensional design permits within-subject comparisons. For example, either stimulus probability, incentive level, stimulus run length, or number of trials preceding a probability reversal could be varied for one dimension and then the generalization effects observed by studying prediction changes on the other (control) dimension. Consequently, the utility of the present design will perhaps be demonstrated when variables found to influence one-dimensional probability learning are studied in the two-dimensional case.

EXPERIMENT II

Method

—Subjects—Fourteen students (seven males, seven females) in the introductory psychology course at Virginia Polytechnic Institute and State University served as subjects.

—Groups—All of the subjects had the same sequence of 250 stimulus presentations in which the probabilities of an up or down (D) *shape*, a red or green (G) *color*, and a left or right (R) *location* were varied on independent schedules. That is, each presentation was a red or green triangle, with its apex up or down, that occurred at the left or right of the stimulus screen. For each successive block of 50 trials, the shape values were presented on a 90%-10% distribution (45 down, 5 up shapes), the color values on a 70%-30% schedule (35 green, 15 red colors), and the location values on a 50%-50% distribution (25 left, 25 right locations). Counterbalancing all possible combinations of the stimulus alternatives for three binary dimensions would have required eight experimental conditions and was beyond the scope of the present study.

—Procedure—Before each three-dimensional presentation, the subjects predicted which shape, color, and location they expected to occur. As in Experiment I, they were given no information about the sequence of dimension alternatives but were told that the stimulus sequence was predetermined by a punched paper tape. Before the 250 experimental trials, 10 practice trials were given, during which the subjects' questions about their task were answered.

Results and discussion

Figure 2 depicts the proportion of *down*, *green*, and *right* predictions for each successive block of 25 trials. Although the subjects' prediction proportions for each dimension did not match each dimension's probabilities, the early separation of the three curves indicates that the subjects did discriminate between the three simultaneously presented probability distributions. The overall analysis of variance (Table 3) indicated significant main effects of dimension and trial block and a significant interaction of dimension and trial block.

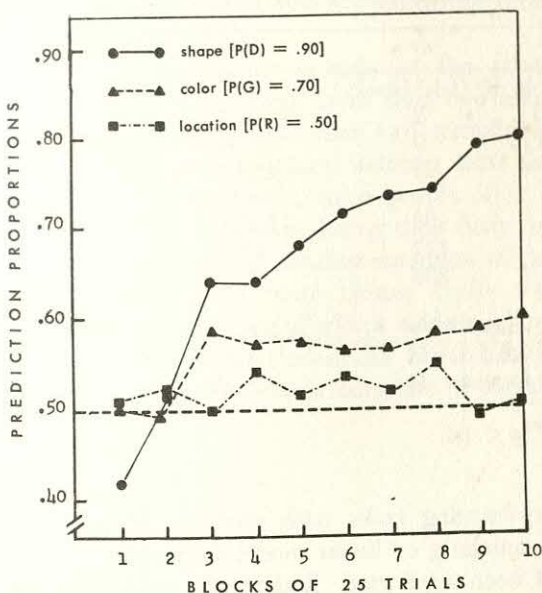


Fig. 2. The proportion of down (D), green (G), and right (R) predictions for each successive block of 25 trials; P is the theoretical probability

As Figure 2 shows, the proportion of downs predicted increased consistently over successive trial blocks and in fact did not appear to reach an asymptotic level over the 250 trials. For the color dimension, the prediction proportions of the more frequent alternative (green) increased successively over the initial trial blocks, and after three blocks (75 trials) apparently reached an asymptotic level slightly below .60. Finally, the proportion of right locations predicted did not vary consistently as a function of trial block.

The separate analyses of variance to determine the significance of

trial-block effects on the response proportions for each dimension indicated a significant main effect of trial block for the shape and color dimensions, but not for location (see Table 3). From the analyses of variance to determine the effect of the last five trial blocks on the response proportions of each dimension, a significant main effect of trial block was obtained only for the shape dimension [$F(4, 117) = 2.71, p < .05$]. Thus, an asymptotic prediction level was reached for the 50%-50% and 70%-30% schedules, but not for the 90%-10% schedule.

Table 3. Analyses of variance, Experiment II

Source	df	MS	Error	F
3 (dimensions) $\times$ 10 (trial blocks)				
Dimension (D)	2	19,493.40	a	12.84**
Trial block (T)	9	403.04	b	4.62**
D $\times$ T	18	925.53	c	9.33**
Error terms				
a. Ss $\times$ D	26	1,519.11		
b. Ss $\times$ T	117	87.34		
c. Ss $\times$ D $\times$ T	234	99.35		
Simple effects				
Trial block for shape	9	2,003.50	b	22.94**
Trial block for color	9	180.06	b	2.06*
Trial block for location	9	70.46	b	.81

* $p < .05$.** $p < .01$.

In probability-learning tasks with multiple stimulus alternatives, the probability-matching or linear model developed by Estes (1959) has usually not been supported. Rather, the subjects typically overpredict the most frequently occurring event and therefore underpredict the less probable alternatives (Gardner, 1957, 1958; Cotton and Rechtschaffen, 1958). With the two-dimensional design of Experiment I, the probability matching predicted by a linear model was demonstrated when the stimulus alternatives of two independent dimensions were varied on identical probability schedules (each 70%-30%). But when the probability distributions for the two dimensions were different (were 70%-30% and 50%-50%), the subjects' response proportions underapproximated the 70% alternative of the 70%-30% dimension. With the three-dimensional design of Experiment II, prediction proportions for a 50%-50% dimension remained at the level hypothesized by a linear model (.50), while

the response proportions for the 70%-30% dimension reached an asymptote that underapproximated the probability of the more frequent alternative. The prediction proportions for the 90%-10% dimension also underapproximated the probability of the more frequent alternative, but since an asymptotic prediction level was not reached for this dimension, the probability learning was not complete. Obviously, the predictive ability of a normative model in a design with multiple dimensions can be realized only after more research with probability schedules and number of trials as parameters. However, the present research did demonstrate that subjects can discriminate between two and three simultaneously occurring probability schedules after relatively few trials.

Since in everyday situations outside of the laboratory, people typically make decisions involving more than two alternatives, probability-learning studies using more than two stimulus alternatives are a more realistic analogue of natural settings than are simple two-alternative tasks. Correspondingly, since people often encounter several probabilistic situations at once during their daily decision making, the present research introduced another extension of the simple probability-learning paradigm, to more closely typify the probabilistic situations of the normal milieu. Perhaps additional research should extend the number of stimulus dimensions beyond three, to make the paradigm an even more appropriate correlate of natural probabilistic settings.

Notes

For their helpful suggestions the authors thank W. B. Pavlik and J. J. Franchina; for their assistance in collecting and analyzing data, J. C. Farris, G. J. McGuire, and R. G. Wiley; also Western Electric Company, for donating electronic supplies. Portions of this research were presented at a meeting of the Eastern Psychological Association (Geller, Whitman, and McGuire, 1971). Requests for reprints should be sent to E. Scott Geller, Department of Psychology, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061. Received for publication June 28, 1971.

1. The large $Ss \times D$ error term in Table 1 indicates relatively inconsistent dimension predictions across subjects. Such a result might perhaps be due to individual differences of attention to the two or three dimensions. However, the small $Ss \times D \times T$ error term suggests that the attentional strategy adopted by individual subjects was quite consistent over trials. This phenomenon clearly requires further investigation in its own right.

References

- Estes, W. K. 1959. Component and pattern models with Markovian interpretations. In *Studies in mathematical learning theory*, ed. R. R. Bush and W. K. Estes, pp. 9-52. Stanford, Calif.: Stanford University Press.
- Estes, W. K. 1964 Probability learning. In *Categories of human learning*, ed. A. W. Melton, pp. 89-128. New York: Academic Press.
- Gardner, R. A. 1957. Probability-learning with two and three choices. *American Journal of Psychology* 70:174-185.
- Gardner, R. A. 1958. Multiple-choice decision-behavior. *American Journal of Psychology* 71:710-717.
- Geller, E. S. 1970. Intrasequential effects of probability and run length on stimulus predictions in a two-stimulus and three-stimulus design. *Proceedings of the 78th Annual Convention of the American Psychological Association* 5:15-16.
- Geller, E. S., Whitman, C. P., and McGuire, G. J. 1971 Probability learning on two simultaneously presented probability schedules. Paper presented to the Eastern Psychological Association, New York.
- Humphreys, L. G. 1939. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *Journal of Experimental Psychology* 25:294-301.
- Schipper, L. M. 1970. Right and wrong feedback in multiple probability learning as a determinant of subsequent performance. *Psychonomic Science* 18:250-251.
- Vlek, C. A. J. 1970. Multiple probability learning: Associating events with their probabilities of occurrence. *Acta Psychologica* 33:207-232.

PERCEIVED COLD AND SKIN TEMPERATURE AS FUNCTIONS OF STIMULATION LEVEL AND DURATION

Lawrence E. Marks and Joseph C. Stevens

John B. Pierce Foundation Laboratory and Yale University

Stimulation was a reduction (ΔH) in the magnitude of heat flux irradiating subjects in cold air. For any constant duration, the terminal cold experienced grew as a power function of ΔH ; the exponent ranged from 1.2 for the longest duration (30.9 sec) to 1.7 for the shortest (3.9 sec). For any constant ΔH , cold, unlike warmth, increased continuously and markedly with duration. Perceived cold at low, but not at high, levels of ΔH correlated strongly with changes in skin temperature.

This study examines how the magnitude of perceived cold (cold sensation) and of skin temperature depend on the length of time the body is exposed to various levels of cold stimulation. In an experiment, a given subject lay supine in very cold air and was irradiated from above by heat lamps, regulated to make him comfortable. From time to time the irradiation was abruptly lowered in order to arouse a sensation of cold, whose magnitude the subject attempted to assess numerically (the method of magnitude estimation).

Two parameters were varied: the magnitude and the duration of the change of irradiation. An analogous experiment on warmth aroused by irradiation of the front of the body (Marks and Stevens, 1968) revealed that warmth grows as a power function of level of irradiance, given any constant duration (between 2 and 12 sec). As a function of duration, however, the results were not so simple. During a brief period (about 15 sec) after onset of stimulation and before the start of vasodilation and sweating, warmth changes little with duration, even though skin temperature rises continuously and markedly. That the sensation of warmth does not clearly parallel skin temperature may mean that a time-dependent process is at work, such as adaptation. The present study sought to learn whether similar processes occur under cold stimulation in a short period (about 30 sec) during which the skin behaves as a thermally passive system, that is, before the onset of vasoconstriction and so on.

The main questions asked here are four. Does cold sensation conform to the psychophysical power equation at various constant durations of exposure? If so, do the constants of the power equation depend on duration? What is the nature of the buildup of cold sensation over time? Are sensations of cold related in a simple and direct manner to skin temperature?

METHOD

—Apparatus—The experiments took place in a large (3.4 by 4.3 by 7.6 m) air-conditioned room. Air temperature was 3–4 deg C, relative humidity below 30%, and average wind velocity approximately .3 m/sec as measured with a Hastings air meter (RM-1X) placed at the level of the subject's chest. Each subject, blindfolded and clad in briefs, socks, and shoes, lay supine on an .8-by-2.0-m bed that was covered with thick foam rubber (see Figure 1). The ventral surface of his body was irradiated from a bank of eight General Electric T-3 quartz lamps, each 40 cm long. The lamps were contained in parabolic, cylindrical reflectors mounted 1.65 m above the bed. The level of radiation emitted by these lamps could be adjusted for each subject individually in order to produce thermal neutrality in the cold air.

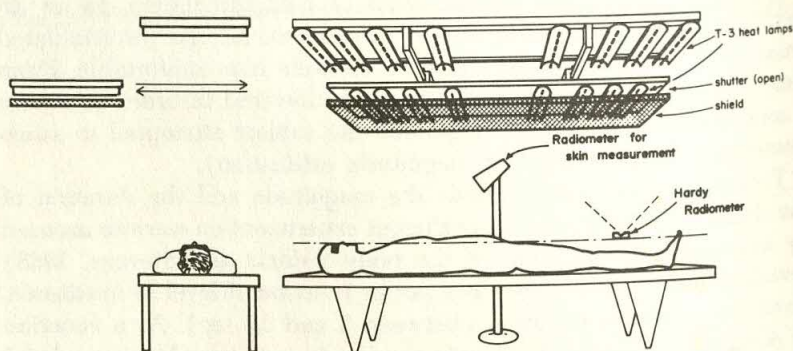


Fig. 1. The experimental setup: at the left an end view; at the right, a side view

Fifteen cm below this bank was another bank of eight heat lamps, mounted on the underside of a shutter that could be moved away quickly by an air cylinder. When the shutter was closed, the subject was exposed to the lower bank; when it was open, to the upper bank. An aluminum shield prevented the lower bank from radiating his body when the shutter was open. This arrangement permitted a relatively abrupt change from one level of irradiation to another, an abrupt change that could not be achieved

by reducing the voltage across a single bank of lamps because it took too long for irradiance to reach a new steady state.

Irradiance from the lamps was controlled by varying a voltage across the lamps. The relation between voltage and emitted radiation was measured with a standardized Hardy radiometer (Hardy, Wolff, and Goodell, 1952) placed at the level of the subject's skin. The lamps were spaced in each bank to produce an irradiant flux that was uniform within 10% over the surface of his body. However, the *change* in irradiance produced by switching from one bank to another varied to a greater extent over the body surface (as much as 25%). For that reason, irradiance was measured at 30 positions in the plane of the subject's body. These positions were selected to represent approximately equal areas of the exposed surface of the trunk and legs of a typical subject. Changes in irradiance reported below are averages for the 30 positions. Because not all of the incident radiation is absorbed by the skin, the measured irradiances were corrected for skin reflection (see Gagge, Rapp, and Hardy, 1967; Stevens and Marks, 1967).

A chopped-beam infrared radiometer was used to measure skin temperature (see Stolwijk and Hardy, 1965). It was mounted above and to the side of the subject so as to 'view' the left central region of his chest.

—Procedure—The purpose of the first experimental session was to determine, for each subject, the approximate level of irradiance from the upper bank of lamps required to produce subjective *thermal neutrality* in the cold air. A staircase procedure was used: 5-V changes were made once every minute. The direction of voltage change depended on the subject's prior response: when he responded "warm," the voltage was decreased; when he said "cool," the voltage was increased. The session lasted about one hour, during which time about 20–25 reversals of response were obtained, and then averaged. The irradiance in mW/cm^2 needed for thermal neutrality varied across subjects from 26 to 49 mW/cm^2 (about an 8-deg-C range of operative temperature). This is in line with earlier measurements of thermal neutrality (Gagge and Stevens, 1968), which show that subjects differ in their choice of thermal environment. A given person, however, tends to choose the same environment fairly consistently from one session to another.

For the *magnitude estimations of cold sensation*, the levels of irradiance found to produce subjective thermal neutrality were used as baselines for the exposures to cold. (These latter sessions began with a brief check to establish that the conditions for neutrality had remained essentially the same; occasionally, small adjustments had to be made.) For each subject the set of stimuli consisted of several different magnitudes of *change* in irradiance (ΔH). The magnitudes of ΔH were obtained by adjusting the voltage to the lower bank of heat lamps. The values of ΔH were identical for all subjects: 9.8, 11.8, 13.8, 17.8, 23.8, and 29.8 mW/cm^2 . (For those subjects who required the smallest amount of heat to provide comfort, the largest ΔH was produced with zero voltage across the lower lamps; this $\Delta H = 29.8 \text{ mW}/\text{cm}^2$ corresponded to lying unheated in air at 3–4 deg C.)

Each of these six levels of ΔH was presented through each of six differ-

ent exposure durations. The shutter, with attached lamps, was massive, and therefore took about a second fully to open or close. The temporal course of the change in irradiance was measured with a photocell. The durations (3.9, 5.9, 8.9, 12.9, 20.9, and 30.9 sec) represent the effective stimulus duration, namely, the integral over time of ΔH divided by the peak ΔH .

Three exposures (not judged) were given at the start of each session to acquaint the subject with the stimulus. Thereafter, one of the intermediate combinations of intensity and duration was presented, and the subject assigned it a number that seemed to him most appropriate to stand for the degree of cold he experienced at the *end* of the exposure. This judgment was not counted. To the 36 subsequent exposures he assigned numbers in proportion to the degree of experienced cold (method of magnitude estimation). The order of presentation was irregular with respect to intensity and duration, and was different in every session. The interval between exposures was variable and dependent on the subject's report that he had returned to a state of subjective thermal neutrality. Intervals typically ranged from about 30 sec after the least intense, shortest exposure to several minutes after the most intense, longest one. Each of 14 subjects took part in two such sessions on different days, judging the same set of stimuli both days.

How *skin temperature* declined with the magnitude and duration of ΔH was determined in subsequent measurements on six of the subjects. Each of the six levels of ΔH was presented for 30 sec, and skin temperature was monitored continuously with the chopped-beam radiometer; intervals of 50 to 300 sec, depending on the preceding level, separated presentations. Each level was presented twice in the session.

RESULTS AND DISCUSSION

Cold as a function of stimulus intensity

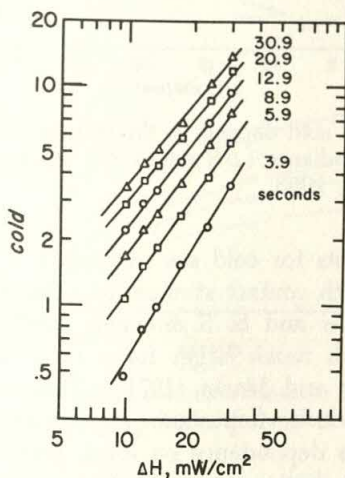
The geometric means of the magnitude estimations of cold are plotted in Figure 2 as functions of the decrease in irradiance (ΔH) from the level that produced thermal neutrality. For each of the six (constant) durations, cold (ψ) turned out to vary as a power function of ΔH , according to the equation

$$\psi = k(\Delta H)^\beta.$$

Sensations of cold produced by changes in radiation of the body conform, therefore, to the power law (S. S. Stevens, 1961). The parameters k and β both depend on duration: the size of k (reflected by the 'intercepts' in Figure 2) increases with duration, but the size of β (the slopes in Figure 2) decreases. Table 1 lists these constants as determined by the method of least squares.

Table 1. Parameters of the equation $\psi = k(\Delta H)\beta$ where ΔH is in mW/cm^2

Duration (sec)	β	k
3.9	1.71	.011
5.9	1.44	.042
8.9	1.34	.078
12.9	1.28	.115
20.9	1.22	.18
30.9	1.18	.23

Fig. 2. Magnitude estimation of cold as a function of the magnitude of the decrease in irradiance (ΔH); the parameter is the duration of stimulation

That β was always larger than 1.0 (range: 1.2 to 1.7) means that cold magnitude grows as a positively accelerated function of ΔH . The corresponding exponents for warmth are smaller; for durations shorter than about 12 sec, warmth grows as a negatively accelerated function of increases in irradiance (Marks and Stevens, 1968). Figure 3 compares directly, in linear coordinates, how cold and warmth depend on stimulus intensity for an exposure duration of 6 sec.

The family of cold functions in Figure 2 tends to converge, on extrapolation, to a common point corresponding to ΔH of roughly $1 \text{ W}/\text{cm}^2$ (a magnitude not physically realizable). In contrast, the analogous family of warmth functions tends slightly to diverge. The longer the duration, the larger the size of the exponent of the power function that relates warmth to increase in irradiance.

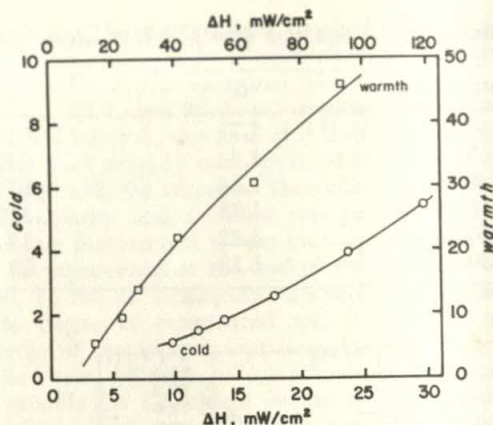


Fig. 3. How warmth and cold depend on the magnitude of the change (increase or decrease) in irradiance (ΔH), for 6-sec exposures; data on warmth from Marks and Stevens, 1968

The present exponents for cold are somewhat larger than the one ($B \approx 1.0$) obtained with contact stimulation of 3.14 cm^2 of the forearm for 3 sec (Stevens and S. S. Stevens, 1960). In contrast, the exponent for warmth is much larger for a small skin area than for a large one (Stevens and Marks, 1971). The dependence of the exponent on area reflects an important property of spatial summation of warmth, namely, its dependence on level. Should it turn out that the exponent for cold does not vary with area, then it is probable that the rules that govern spatial summation of warmth and of cold differ considerably. Much remains to be learned, however, about how cold depends on the locus and areal extent of stimulation.

Cold as a function of stimulus duration

Figure 4 shows the estimates of cold, here as functions of duration. At all levels of ΔH , the magnitude of cold increased regularly and markedly with time. Cold grows as a positively accelerated function of stimulus level (see Figure 3), but as a negatively accelerated function of duration. The *relative* size of the increase in cold magnitude over time changed with stimulus level (a fact implicit in the convergence of the family in Figure 2). For example, at the lowest level, cold increased by more than 600% between 3.9 and 30.9 sec, but at the highest level by less than 300%. The *absolute* increase in cold was, however, much larger at the highest stimulus level. Negatively accel-

erated time curves have also been proposed by Bujas (1939), who obtained various combinations of intensity (cooled air blown on the forehead) and duration that matched for apparent cold. (Bujas's results suggested, however, that cold sensation ceases to grow after about 8 sec.)

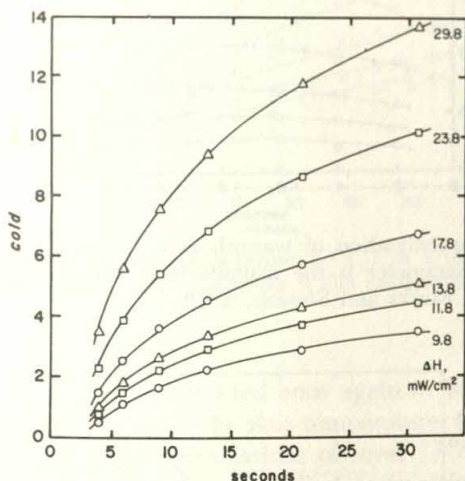


Fig. 4. Magnitude estimation of cold as a function of the duration of stimulation; the parameter is the magnitude of the decrease in irradiance (ΔH)

The time course for warmth is quite different (see Figure 5). At the two highest stimulus levels presented, warmth increased modestly with duration. Lower levels gave little or no increase in warmth, and there was a suggestion of a small rise and subsequent fall at the lowest levels.

Relation of cold to skin temperature

Changes (decreases) in skin temperature (ΔT_s) were read from the continuous records at 3-sec intervals. Arithmetic averages of the changes at six of the intervals are plotted in Figure 6 as functions of ΔH . (Averages at other intervals were omitted to improve clarity.) The points can be fitted fairly well by straight lines that pass through the origin. This outcome supports the contention that the thermal behavior of the skin was passive during stimulus exposure: it is not likely that blood flow changed significantly during stimulation at any intensity, for example.

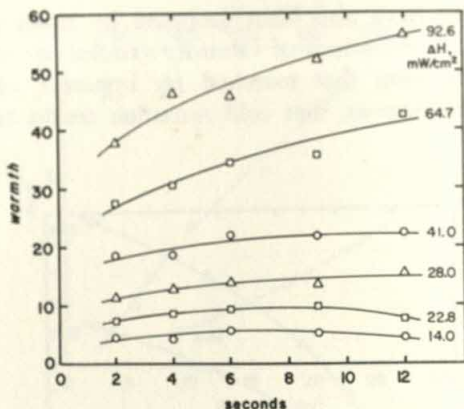


Fig. 5. Magnitude estimation of warmth as a function of the duration of stimulation; the parameter is the magnitude of the increase in irradiance (ΔH); data from Marks and Stevens, 1968

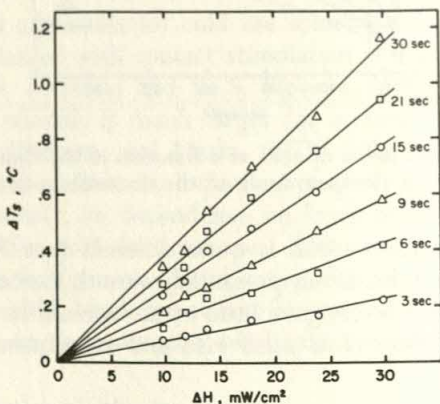


Fig. 6. The decrease of skin temperature (ΔT_s) as a function of the magnitude of the decrease in irradiance (ΔH); the parameter is the time after onset of stimulation

The averages of ΔT_s are plotted in Figure 7 as functions of duration. The curves have been drawn to pass through the points predicted by the straight lines in Figure 6. The shapes of the curves superficially resemble those that relate cold to duration (Figure 4), but unlike the rise of subjective cold with time, the changes in skin temperature are proportionally the same at all levels of stimulus intensity.

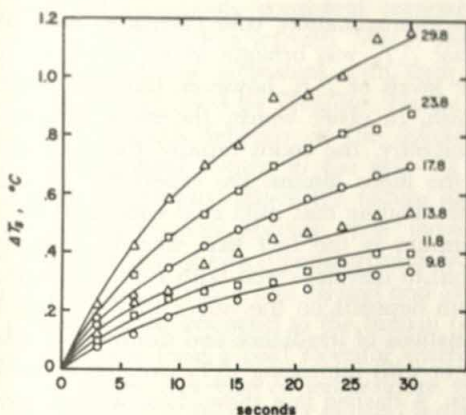


Fig. 7. The decrease of skin temperature (ΔT_s) as a function of the time after onset of stimulation; the parameter is the magnitude of the change in irradiance (ΔH)

The estimates of cold are plotted once again in Figure 8, this time as functions of the decreases in skin temperature (ΔT_s) that were produced by the various combinations of level (ΔH) and duration. For small values of ΔH , the results fall close to a single curve, mean-

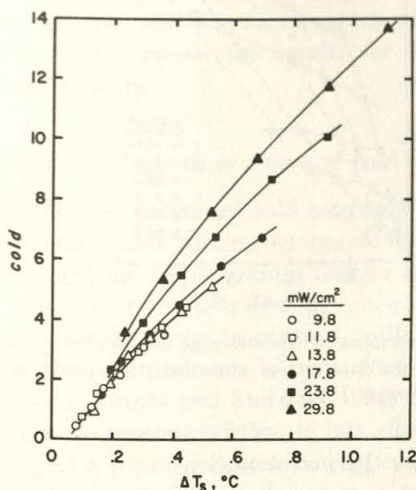


Fig. 8. How cold relates to the decrease of skin temperature (ΔT_s) produced by different combinations of stimulus magnitude and duration

ing that to a first approximation, cold correlates well with ΔT_s , regardless of whether ΔT_s was brought about slowly or quickly. With higher and higher levels of ΔH , however, the correlation tends more and more to vanish. In other words, the more intense the stimulation, and concomitantly, the more rapidly the skin is cooled by a certain amount, the more intense the experience of cold.

In contrast to the finding that mild cold correlates well with superficial skin temperature is the fact that warmth sensations, whether intense or mild, fail to correlate well with skin temperature. Figure 9 shows how warmth depends on the values of ΔT_s that are produced by various combinations of irradiance and duration. A solid line shows how warmth grows with increasing ΔT_s over time under a constant level of irradiation. A dashed line shows how warmth grows with the increasing ΔT_s that is produced by increasing the level of constant-duration stimulation. Change in skin temperature alone does not predict warmth magnitudes at any level of stimulation. At all levels of irradiance, the more rapidly the skin is heated by a certain amount, the more intense is the warmth sensation.

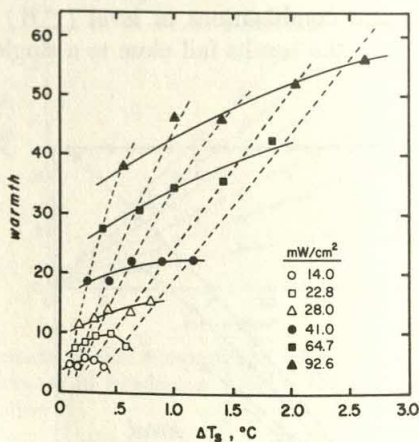


Fig. 9. How warmth relates to the increase of skin temperature (ΔT_s) produced by different combinations of stimulus magnitude and duration; data from Marks and Stevens, 1968

Proximal stimulus for thermal sensation

Both cold and warmth depend in complex ways on the intensity and duration of radiant stimulation. Furthermore, induced changes in

skin temperature fail to provide consistent correlates to the magnitudes of thermal sensations. For warmth, Marks and Stevens (1968) found two hypotheses to be consistent with their results. The first hypothesis, originally suggested by Hendler, Hardy, and Murgatroyd (1963), states that warmth depends on the difference between the temperature at two critical depths in the skin—at approximately .2 and 1.0 mm below the surface. Since the nature and location of the receptors for warmth remain obscure, this hypothesis remains a possibility. But such a hypothesis appears extremely unlikely for cold, if Hensel and Zotterman's (1951) and Hensel and Witt's (1959) results for the cat's tongue can be extended to the human perception of cold. They obtained responses from a cold receptor when the tongue was cooled, regardless of the direction of temperature gradient.

The second hypothesis of Marks and Stevens involved an appeal to adaptation in the warmth system. Temperature changes produced by rapid heating yield greater warmth than identical changes produced by slower heating. That result is consistent with the assumption that continuous adaptation serves to offset the effect of continuously increasing skin temperature, for rapid heating would provide less opportunity for adaptational processes to come into play. A similar hypothesis could help to account for the way cold depends on skin temperature. However, Figure 8 makes it clear that this hypothesis would be required only for the highest levels of intensity, when the initial rate of skin cooling is relatively large. Any hypothesis about the relation between cold and skin temperature will have to take into account the difference between the results for low and for higher levels of stimulus intensity.

Functional significance of the time course of cold sensation

Even at low intensities, perceived cold increases steadily as exposure continues. This marked and continuous rise in the sensation of cold may provide early signals of impending danger to the well-being of the organism. With long-duration cooling, mean body temperature falls markedly and continuously; in contrast, with long-duration heating, mean body temperature rises slowly, then asymptotes within about 30 min or less (Hardy and Stolwijk, 1966; Stolwijk and Hardy, 1966). This is because vasoconstriction is less effective than vasodilation and sweating as a means for maintaining constant internal body temperature. Continuous cooling is therefore a greater threat to the organism than is continuous heating. This difference appears to have a

parallel in the different time courses for the sensations of cold and warmth.

The discomfort produced by thermal stimulation also appears to obey different rules depending on whether the discomfort is produced by cooling or by heating (Stevens, Marks, and Gagge, 1969). The exponent of the power function that describes the rate of growth of discomfort produced by cooling the skin (1.7) is more than twice as large as the exponent (.7) that describes discomfort produced by heating the skin. The larger exponent for cold discomfort may signify the greater threat posed by body cooling.

Notes

This research was supported by Grant AFOSR-70-1950 with the Air Force Office of Scientific Research. Requests for reprints should be addressed to Dr. L. E. Marks, John B. Pierce Foundation, 290 Congress Avenue, New Haven, Connecticut 06519. Received for publication July 14, 1971.

References

- Bujas, Z. 1939. L'évolution de la sensation du froid en fonction de la durée d'excitation. *L'Année Psychologique* 39:184-198.
- Gagge, A. P., Rapp, G. M., and Hardy, J. D. 1967. Effective radiant field and operative temperature necessary for comfort with radiant heating. *Transactions of the American Society of Heating, Refrigerating, and Air-conditioning Engineers* 73 (part 1).
- Gagge, A. P., and Stevens, J. C. 1968. Thermal sensitivity and comfort. In *The skin senses*, ed. D. R. Kenshalo, pp. 345-367. Springfield, Ill.: Charles C Thomas.
- Hardy, J. D., and Stolwijk, J. A. J. 1966. Partitional calorimetric studies of man during exposure to thermal transients. *Journal of Applied Physiology* 21:1799-1806.
- Hardy, J. D., Wolff, H. G., and Goodell, H. 1952. *Pain sensations and reactions*. Baltimore: Williams and Wilkins.
- Hendler, E., Hardy, J. D., and Murgatroyd, D. 1963. Skin heating and temperature sensation produced by infrared and microwave irradiation. In *Temperature: Its measurement and control in science and industry*, vol. 3, part 3, ed. C. M. Herzfeld, pp. 211-230. New York: Reinhold.
- Hensel, H., and Witt, I. 1959. Spatial temperature gradient and thermoreceptor stimulation. *Journal of Physiology* 148:180-189.
- Hensel, H., and Zotterman, Y. 1951. Action potentials of cold fibers and intracutaneous temperature gradient. *Journal of Neurophysiology* 14:377-385.
- Marks, L. E., and Stevens, J. C. 1968. Perceived warmth and skin temperature as functions of the duration and level of thermal irradiation. *Perception and Psychophysics* 4:220-228.

- Stevens, J. C., and Marks, L. E. 1967. Apparent warmth as a function of thermal irradiation. *Perception and Psychophysics* 2:613-619.
- Stevens, J. C., and Marks, L. E. 1971. Spatial summation and the dynamics of warmth sensation. *Perception and Psychophysics* 9:291-298.
- Stevens, J. C., Marks, L. E., and Gagge, A. P. 1969. The quantitative assessment of thermal discomfort. *Environmental Research* 2:149-165.
- Stevens, J. C., and Stevens, S. S. 1960. Warmth and cold: Dynamics of sensory intensity. *Journal of Experimental Psychology* 60:183-192.
- Stevens, S. S. 1961. The psychophysics of sensory function. In *Sensory communication*, ed. W. A. Rosenblith, pp. 1-33. New York: Wiley.
- Stolwijk, J. A. J., and Hardy, J. D. 1965. Skin and subcutaneous temperature changes during exposure to intense thermal radiation. *Journal of Applied Physiology* 20:1006-1013.
- Stolwijk, J. A. J., and Hardy, J. D. 1966. Partitional calorimetric studies of responses of man to thermal transients. *Journal of Applied Physiology* 21:967-977.

ANNOUNCEMENT

NSF Chautauqua-type Short Courses for College Teachers

The American Association for the Advancement of Science has announced that the third annual program of NSF Chautauqua-type Short Courses will begin in October 1972. There will be places for 3,000 participants, with ten courses being offered at each of 12 Field Centers located throughout the country.

Topics cover a wide range of fields in the natural and social sciences, mathematics, and engineering. Most of the courses are interdisciplinary in nature, including the applications of mathematics to college teaching.

The courses consist of two days of intensive classwork at one of the centers, followed by approximately three months of independent research or study, and ending with another two-day session for discussion of the work performed and general synopsis.

The announcement brochure of the 1972-1973 program can be obtained by writing to NSF C-T Courses, AAAS, 1515 Massachusetts Avenue, N.W., Washington, D.C. 20005.

IMPLICIT CONTOURS IN THE ZÖLLNER ILLUSION

K. G. White

University of Otago, Dunedin, New Zealand

For Zöllner illusion figures with dotted background lines, the magnitude of illusion systematically decreased as dot separation increased. It is suggested that the dotted lines function as implicit contours and that the perceptual enlargement of the implicitly defined acute angles contributes to the overall illusion.

Many of the classical geometrical illusions consist of a number of lines intersecting at acute angles. The spatial distortion occurring in these figures is primarily a function of the angle size, and it appears that as the number of angles increases so does the overall illusion (Piaget, 1969; Wallace and Crampin, 1969). That is, the illusion can be considered as the sum of effects of the individual angles. In general, the magnitude of a single acute angle is overestimated, and in this respect the geometrical illusions can be regarded, at a descriptive level, as instances of the perceptual enlargement of acute angles (Berliner and Berliner, 1948; Day, 1965; Wallace, 1966).

It would seem then, that a satisfactory explanation for the illusions should mainly concern itself with the contrast effect produced by a single acute angle. Does this require specifically defining an angle in terms of the intersection of *explicit* lines? Would acute-angle enlargement still be a useful descriptive concept if it were found that the illusion persists for implicit contours?

The present experiment examined the relationship between line 'density' and the magnitude of the Zöllner illusion for figures with dotted inducing or background lines. The choice of the Zöllner illusion was suggested by its relationship to acute-angle enlargement (White, 1971).

METHOD

—Stimuli—Each of the four stimulus figures (see Figure 1) comprised a series of background lines spaced 15 mm apart and inclined at 15 deg to the horizontal. In one instance the 1-mm-wide background lines were con-

tinuous. Otherwise, the background lines comprised rows of dots 1 mm in diameter and placed either 5, 10, or 15 mm apart. The circular stimulus field was 15.5 cm in diameter. Two 1-mm-wide test lines were superimposed on the background and could be varied in tilt. The test lines were 40 mm apart at the fixed left-hand end, and adjustment of a wheel by the subject could effect their simultaneous convergence or divergence. The luminance of the figure was approximately 35 mL.

—Subjects and procedure—Eight male and two female subjects with a mean age of 20.5 years served. The seated subject monocularly viewed the figure placed in the frontoparallel plane at a distance of 60 cm. Before each judgment the test lines were obviously divergent, and the subject was instructed to position the two lines so that they looked as though they were parallel. After each judgment the test lines were returned to the divergent position and a new background pattern was inserted. Each subject made five adjustments for each of the four figures. In addition, estimates of the zero error were obtained from five judgments made when the test lines were superimposed on a blank background. In order to balance the effects of practice over time, presentation of the background patterns was randomized over blocks of five. Different random orders were used for each subject. The magnitude of illusion was measured in terms of the angular deviation of the upper test line from the horizontal. The horizontal was in turn defined for each subject as the zero error.

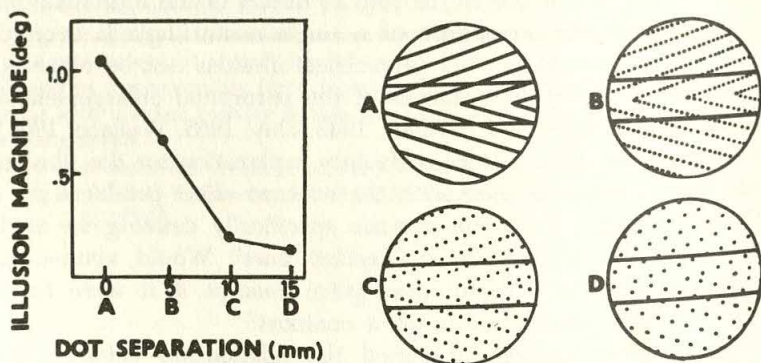


Fig. 1. Magnitude of illusion as a function of dot separation for the Zöllner illusion drawn with continuous background lines (A) or with dotted lines (B, C, D)

RESULTS AND DISCUSSION

For each subject the mean zero error was subtracted from the mean estimate of parallel obtained for each illusion figure. Analysis of variance revealed that the effect of dot separation [$F(3, 27) = 142.98$, $p < .001$] was highly significant. Figure 1 shows mean illusion mag-

nitude as a function of dot separation. As the dot separation increased the illusion's magnitude systematically decreased.

The present result supports the notion that dotted lines function as implicit contours. In this respect, the stimulus figures used in the present experiment can be said to contain acute angles, the perceptual enlargement of which contributes to the overall illusion.

Geometrically, a straight line can be considered as a row of dots with zero separation. The criterion for when a line is not a line is given by a certain dot separation, dependent on the nature of surrounding implicit contours. In the present instance the surrounding contours were virtually unrecognizable as such when the dot separation was 15 mm, that is, when it was equal to the perpendicular distance between adjacent contours. An implicit line can also be defined between two single dots. For thus-dotted Müller-Lyer and Poggendorff figures, the illusion is reduced (Coren, 1970). Dotting a line results in a decreased overall intensity of the corresponding explicit contour and a diminished illusion, perhaps in the same way that contrast reduction results in reduced illusion magnitude (Wickelgren, 1965). Similarly, the contrast effect of an acute angle decreases with decreasing dot density of the implicit contours defining it.

Ganz (1966) attributes interaction between spatially adjacent contours to lateral-inhibitory processes in the visual system. Despite the inability of the lateral-inhibition model to account for the effects of judgmental variables on illusion magnitude (Over, 1969), the model directly predicts the contrast effect in acute angles as well as diminished illusion with a reduction in contour intensity. Requisite to the production of lateral inhibition, however, is the presence of *explicitly* defined contours. Dots do not seem to generate inhibition (Békésy, 1968), and the inhibitory model cannot thus be applied to the interaction between implicit contours. Yet, the evidence for feature analyzers in the visual system suggests that a model for the geometrical illusions which utilizes the inhibitory and excitatory properties of orientation and angle detectors in the human visual cortex is not beyond consideration—especially if the information provided by implicit contours is found to be equivalent to information generated by low-intensity explicit contours.

Notes

Received for publication June 28, 1971.

References

- Békésy, G. von. 1968. Mach- and Hering-type lateral inhibition in vision. *Vision Research* 8:1483-1499.
- Berliner, A., and Berliner, S. 1948. The distortion of straight and curved lines in geometrical fields. *American Journal of Psychology* 61:153-166.
- Coren, S. 1970. Lateral inhibition and geometric illusions. *Quarterly Journal of Experimental Psychology* 22:274-278.
- Day, R. H. 1965. Inappropriate constancy explanation of spatial distortions. *Nature* 207:891-893.
- Ganz, L. 1966. Mechanism of the figural aftereffect. *Psychological Review* 73:128-150.
- Over, R. 1969. Lateral inhibition explanation of geometrical illusions. *Nature* 222:99-100.
- Piaget, J. 1969. *The mechanisms of perception*, trans. G. N. Seagram, pp. 14-49. New York: Basic Books.
- Wallace, G. K. 1966. Optical illusions. *Nature* 209:327-328.
- Wallace, G. K., and Crampin, D. J. 1969. The effect of background density on the Zöllner illusion. *Vision Research* 9:167-177.
- White, K. G. 1971. Zöllner illusion as perceptual enlargement of acute angle. *Psychonomic Science* 24(2):73-75.
- Wickelgren, B. G. 1965. Brightness contrast and length perception in the Müller-Lyer illusion. *Vision Research* 5:141-150.

ORTHOGRAPHIC DISTINCTIVENESS AS A VARIABLE IN WORD RECOGNITION

Eugene B. Zechmeister

Loyola University, Chicago

Of low-frequency words scaled for orthographic distinctiveness, the structurally interesting or unusual (high distinctiveness) words were recognized better, and were chosen less often when used as distracters, than words rated not distinctive in orthography. Confidence judgments in correct recognition responses were significantly affected by the presence of a distinctive test distracter. Orthographic distinctiveness may be encoded as part of the memory for a word.

There is compelling evidence that a frequency attribute dominates recognition-memory performance (Underwood, 1969; Underwood and Freund, 1970). On the other hand, it is obvious that frequency may not be the only attribute by which discrimination between 'old' and 'new' test words is accomplished. Depending on the nature of the task material and retention test, other attributes may serve to facilitate discrimination or to contribute additional cues by which one may judge the correctness of the recognition decisions he has made. For example, as an extreme case, if the study items are of one class of words (say, nouns) and the items used as distracters represent another word class (say, adjectives), this word characteristic could influence one's recognition performance over and above that allowed by differential situational frequency.

A recently investigated characteristic of verbal items is orthographic distinctiveness (Zechmeister, 1969). Distinctive orthography is defined as those structural characteristics of a word that make it physically unusual or interesting. Correlates of orthographic distinctiveness include low bigram totals, instances of rare letter combinations, and variant typography. For example, of 150 five-letter, low-frequency words rated for distinctive orthography, the words xylem, sylph, and phlox were rated as very distinctive, whereas the words parse, scone, and poser were not.

The present study considered whether distinctive orthography would lead to differential recognition performance when high- and low-

distinctiveness words served as both study and distracter items. That is, it was assumed that words of very distinctive orthography would be recognized better, and be less likely chosen distracters, than words rated as not distinctive. Furthermore, confidence ratings were taken to investigate whether orthographic distinctiveness is related to expressed certainty in correct recognition judgments.

METHOD

—Lists—All items were chosen from 150 low-frequency words (appearing 12 times or less per 18 million according to Thorndike and Lorge, 1944) that had previously been scaled for distinctiveness of orthography (Zechmeister, 1969). List construction was done this way. From the 150 words rank-ordered according to mean distinctiveness ratings, every other word was selected to make two sublists of 75 words. If the original list is considered as representing five arbitrary levels of distinctiveness (30 words in each of five levels in order of mean ratings), then the two sublists had 15 words in each of these five levels. The original 30-word groups in order of high to low orthographic ratings and their corresponding ranges were: group I, 7.09–5.54; group II, 5.53–4.97; group III, 4.94–4.58; group IV, 4.57–4.14; and group V, 4.13–3.06.

For each sublist, 12 words were randomly taken from both the I and V groupings to serve as study items, and 36 items from levels II, III and IV were randomly selected (12 each) to serve as filler items. The words were assigned to the *study lists* such that filler items occupied the first six and last six places in each list; the 24 test items and the remaining 24 filler items were randomized with the restriction that six high- and six low-distinctiveness words appear in each of the two within-list blocks of 24 study items. No more than two high- or two low-distinctiveness study words appeared consecutively. Therefore, two different 60-item study lists were constructed, and each was presented in forward and reverse order to different subjects.

Construction of the *test lists* involved random pairing of the 12 high- and 12 low-distinctiveness study items from the two study lists to achieve four types of study/distracter test pairs (high-high, high-low, low-high, low-low). Two different 24-pair test lists were derived in such a way that no two words of a test pair had the same initial letter and that no two pairs were the same between test lists. Words in the test pair were presented one above the other, with half of each of the four types of test pairs having the study item in the top position, and half in the bottom position. Ordering of the test lists was done so that the 24 test pairs (six each of the above combinations of study and distracter words) represented two blocks of 12 items, each of these test blocks measuring responses to half of the first and half of the second study-block items under (equal numbers of) each of the four test conditions. No words were tested consecutively that appeared in the study list contiguously.

Therefore, eight independent test groups were defined by factorially com-

bining the two study lists with two orders of study-list presentation (forward or backward) and two test lists. Within each group, the subjects were tested for recognition from 24 pairs of old and new items, each representing one of the four study/distracter conditions detailed above.

—Procedure—A slide projector was used to present both study and test lists to small groups of subjects. Presentation rate was 2 sec, with test pairs shown for 8 sec. The subjects were told that their memory for words would be tested and that they would see a long list of words, each shown singly for a brief time. They were advised that the words were uncommon English words and that it was not important if they did not know the meaning for many of them. Before the recognition test, they were given a sheet numbered from 1 through 24 with the letters T and B in the first two columns for indicating the top or bottom word in the slide pair as the one previously seen. Also, next to the space for the recognition response was a five-point scale for marking confidence in each recognition choice. All subjects were told to work quickly and to circle both a letter and a number for each test pair. Time between study and test procedures was approximately 3 min.

—Subjects—The eight conditions represented by the factorial arrangement of the between-group variables (study lists, study order, and test lists) were randomly assigned to time slots and the subjects signed up in small groups for the experiment according to the various times available. All of the subjects were introductory psychology students. At least eight of them were in each test session; additional subjects were randomly dropped to provide equal Ns of eight per test group.

RESULTS

The overall design was a 2^5 mixed factorial with within-subject comparisons on word-distinctiveness conditions. However, since list and order effects add no information for an understanding of the present findings, between-group effects will not be considered further. Also, since recognition involved a forced choice from paired alternatives, scores for any individual could vary from 0 to 6, with chance, of course, .50. Therefore, most scores ranged from 3 to 6 and an arcsine transformation was made of the number of correct recognition responses before an analysis of variance.

Figure 1 shows the percentage of errors made in each of the four conditions achieved by combining high- and low-distinctiveness study items with high- and low-distinctiveness distracter words. Total number of errors made by all 64 subjects was 189, or only 12.30% of the possible recognition choices. Analysis of the number of correct recognition responses showed that both study and distracter words were

significant as main effects [$F(1, 56) = 5.95, p < .05$ for study items, and $F(1, 56) = 10.33, p < .01$, for distracter items]. The interaction between study and distracter items was not significant ($F = 1.15$), although Figure 1 suggests a tendency in that direction. Since errors for the pairs containing two high-distinctiveness words were very few overall (they represented only 15.31% of the total errors made), this apparent ceiling on number correct may be partly responsible for the failure to obtain an interaction.

Mean confidence ratings of correct choices for the four types of test alternatives were 4.31 and 4.20 for high-distinctiveness study words paired with high- and low-distinctiveness distracter items, and 4.30 and 4.11 for low-distinctiveness study words with high- and low-distinctiveness distracter items. The effect for orthographic distinctiveness of distracter items was highly reliable [$F(1, 56) = 11.62, p < .01$]. Subjects were significantly more confident of their recognition decision when the distracter was a structurally unusual word. The effect for distinctiveness of study items was not significant, nor was the interaction.

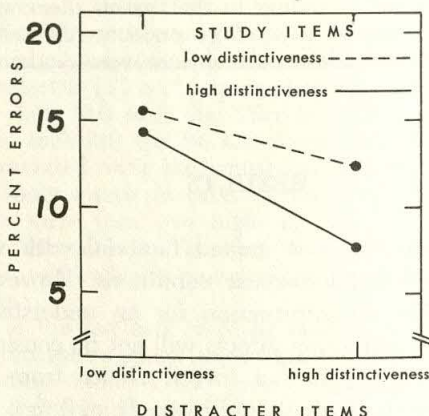


Fig. 1. Percentage of recognition errors for each of four conditions achieved by combining high- and low-distinctiveness study and distracter items

Errors increased minimally as a function of testing. The increase in errors across all conditions was only .67% ($F < 1$) for second-half test performance over that for the first 12 test pairs. However, percentage of errors was significantly less for items from the first study block than for those from the second block of study words (10.41%

versus 14.19%; $F = 7.78, p < .01$). Although the number of errors was small, this effect suggests that items appearing late on the study list were less well encoded than items first presented. Both the homogeneity of the study items and the rather rapid presentation rate may have contributed to this deterioration.

DISCUSSION

When recognition is required for low-frequency words, performance is remarkably high, and significantly superior to that obtained when high-frequency words are used (Schulman, 1967; Shepard, 1967; Underwood and Freund, 1970). In part, better recognition performance for the low-frequency items has been attributed to the reduced opportunity for implicit associative responses to occur to unfamiliar words and hence affect frequency of study and test alternatives (Underwood and Freund, 1970). It has also been suggested that low-frequency words are easier to recognize because of unusual structure (Schulman, 1967) or simply because they are 'strange' (Shepard, 1967). When only low-frequency items are considered, the present results confirm the notion that the physical appearance of a word may contribute to differential recognition performance. Furthermore, since a mixed list was used, employing several levels of rated distinctiveness, it is unlikely that the present findings reflect a contrast effect during study. However, the reliable effect for orthographic distinctiveness of distracter items, both in terms of reduced recognition errors and increased confidence, demonstrates that the presence of a structurally unusual word during testing aids discrimination. When given a choice, a subject may be influenced more by what he is sure the correct word is *not* than by the characteristics of the 'old' item.

It seems reasonable to suggest that 'distinctiveness' may be a memory attribute encoded with the target experience. For example, it is not uncommon to hear someone remark that he has forgotten a certain person's name but that he is sure it is an 'unusual' one. By this reasoning, the implicit response of 'strange,' 'never seen that before,' 'unusual,' or such may be elicited by orthographic characteristics and encoded and stored as part of the memory for that event. A concomitant cue in this situation may be difficulty with pronunciation, since distinctiveness of orthography is inversely related to pronunciability (Zechmeister, 1969).

Notes

Thanks are due to Jack McKillip for assistance in both data collection and analysis. Received for publication June 18, 1971; revision, August 11, 1971.

References

- Schulman, A. I. 1967. Word length and rarity in recognition memory. *Psychonomic Science* 9:211-212.
- Shepard, R. N. 1967. Recognition memory for words, sentences, and pictures. *Journal of Verbal Learning and Verbal Behavior* 6:156-163.
- Underwood, B. J. 1969. Attributes of memory. *Psychological Review* 76:559-573.
- Underwood, B. J., and Freund, J. S. 1970. Word frequency and short-term recognition memory. *American Journal of Psychology* 83:343-351.
- Thorndike, E. L., and Lorge, I. 1944. *The teacher's word book of 30,000 words*. New York: Columbia University Teachers College.
- Zechmeister, E. B. Orthographic distinctiveness. 1969. *Journal of Verbal Learning and Verbal Behavior* 8:754-761.

THE EFFECTS OF CONTEXTUAL CHANGES ON COMPONENT RECOGNITION

Frank DaPolito, Douglas Barker, and Joseph Wiant
University of Dayton

Triplets of verbal items were followed by tests for recognition of one of the items, the other two items (context) being intact, permuted, removed from display, or replaced by new items. Recognition was best when the test context was identical to that during training. The data seem to reflect specific retrieval processes, especially the generating of codes from the test item and the comparing of those codes against multiple representations in memory.

A fundamental variable in the retrieval of items from memory is the completeness with which the original pattern of stimulation is presented during the tests. McGeoch (1942) has reviewed a number of experiments which demonstrate that recall performance clearly depends on the degree to which the training stimuli are reinstated on the tests. Similarly, experiments involving alterations in 'incidental' background stimuli have produced substantial reductions in the recall and relearning of paired associates (Peterson and Peterson, 1957; Weiss and Margoulis, 1954). Moreover, experiments using a free-recall paradigm have shown that the introduction of 'cue' words during the test phase may either facilitate or inhibit the recall of individual words (Postman, Adams, and Phillips, 1955; Tulving and Pearlstone, 1966; Bower, Clark, Lesgold, and Winzenz, 1969).

Although the evidence about the effects of altered test conditions on recall and relearning is quite substantial, experimentation with verbal materials and *other* measures of retention, particularly recognition, has only recently been reported (Light and Carter-Sobell, 1970; Tulving and Thomson, 1971). The present experiment was an attempt to extend these recent findings by examining how contextual changes might produce variations in the correct recognition of verbal 'units' previously presented in a compound display.

METHOD

—Subjects—The subjects were 60 introductory psychology students who

participated in the experiment to fulfill a course requirement.

—Apparatus and materials—A two-channel tachistoscope (Scientific Prototype 800F) was used to present the training and testing stimuli. The training materials consisted of three decks of white 5-by-8-in. cards. In all, there were 21 cards per deck, each card containing a triplet composed of three three-letter units typed in black; each letter was 3 cm wide and 5 cm high. The triplets were centered in the visual field so as to form a 3-by-3 matrix of letters.

Each card in deck 1 contained three common English three-letter *words* selected at random from Underwood and Schulz (1960, Appendix D). Each of the cards in deck 2 contained three CVCs taken from the same source (Appendix A) and randomly selected from those having Glaze association values between 60% and 88%. Similarly, each card in deck 3 consisted of three CCCs, also from the same source (Appendix B) and randomly selected from those with association values between 63% and 88%. In all decks the triplets were constructed with the restriction that no triplet contain three items with a common first letter. Finally, there were two replications of each set of materials.

—Design—In essentials, the design employed two exposure durations (1 sec and .1 sec), two replications of materials, three classes of stimulus materials (words, CVCs, and CCCs), and nine test conditions. The variable of exposure duration involved between-subject comparisons, whereas stimulus class and test condition were within-subject comparisons. Since the two replications yielded essentially identical data, the results were collapsed over this variable for subsequent data analysis.

For a given subject, the session consisted of three successive study/test cycles, so that he was tested on all three sets of materials. Of the possible six orders of presenting the materials, one of these was randomly chosen for each subject.

During the study period, the triplets were exposed for either 1 sec (30 subjects) or .1 sec (30 subjects). The order of presenting the cards was permuted for each subject. Also, the two replications of stimulus materials were counterbalanced for exposure durations.

Apart from offering opportunities for contextual manipulations, the stimuli were presented in triplets to increase the likelihood that a subject would rehearse the triplet's units together during the 5-sec intercard intervals. The tenfold difference in exposure duration was an attempt to vary the amount of feature sampling and subsequent storage during the intercard intervals (pilot data showed no differences between .5- and 1-sec exposures). More important, the procedure required attention to each unit (word, CVC, or CCC) of the triplet, since the subject could not know in advance which of the three units would be tested. This obviates, in part, the problem of subjects' using variable self-instructions, which is very likely when incidental (uninstructed) background cues are used, as in previous studies of context. Also, of course, the context was comprised of units in the same class of materials as those to be tested, rather than being the non-verbal units (such as colors or shapes) that are commonly used in paired-associate experiments. Thus, problems related to differential attention and

difficulty between the probed item (target) and contextual cues were minimized.

On test blocks, the subjects viewed triplets (21 cards) and single-unit stimuli (6 cards), each card being exposed for 5 sec. Each test card contained a large (2 cm) black dot that was placed approximately 10 cm to the right of the target unit. The subject gave a recognition response (yes, no) and a confidence judgment (1, 2, 3, or 4) for each marked item. No informative feedback was provided during test cycles, and the order of presenting the cards was rerandomized for every five subjects.

Of the nine conditions of testing, five involved tests of 'old' items presented during the study period. More specifically, a target item (word, CVC, or CCC) was tested under one of five conditions. In the old-context condition, all of the items of the original triplet were presented. In the mixed-context condition, one of the untested items of the test triplet was old and the other was 'new.' (The two possible ways this could happen were balanced over replications. Also, the old context item and the target were in their original serial positions.) In the new-context condition, the target item was tested with both of the original context items being replaced by new items from the same stimulus population. In the permuted-context condition, the order of the original context items was permuted in the test display. In the removed-context condition, the target item was tested in isolation (a single-unit stimulus—just one word, CVC, or CCC).

The remaining four test conditions involved tests of 'new' items not previously presented in the experiment but drawn from the same stimulus population. The new items were tested in four context conditions. In the old-context condition, the new item was marked, and both of the untested items were taken from a previously studied triplet and retained in their original serial position. In the mixed-context condition, one of the untested items in the display was old and the other new, with the same restrictions as the mixed-context condition above. In the new-context condition, all three items in the display were new. In the removed-context condition, the new item (word, CVC, or CCC) was tested in isolation.

On a given test cycle, there were three replications of each of the nine test conditions. Also, in each test deck the serial position of the target item was totally counterbalanced within context conditions.

—Procedure—The subjects were run individually in a single session, which took approximately 25 min to complete. These instructions were read to each subject:

This is an experiment in recognizing what you have already seen. You will first be shown three sets of cards, with three items on each card. When you are ready, a light will expose each card to you for a short time. Try to remember what was on each card. After you have seen all the cards, you will be shown more cards with a dot next to one of the items. Or you will be shown a card with a single item on it. If you recognize this marked item or the single item as one you have seen before, answer "yes" on the answer sheet. If you do not recognize the marked or single item as one you have just seen before, answer "no" on the answer sheet. Next to each

answer space are four boxes numbered 1 to 4. For each answer you are to evaluate your certainty about the yes/no answer you have given. If you were "sure" that you had either seen or not seen the target item previously, circle box 1. If you were "pretty sure," circle box 2. If you doubted that your answer was right, circle box 3. If you were guessing, circle box 4. You will therefore mark a yes/no answer and a certainty estimation for each test card presented to you. Do you have any questions?

Next, the subjects got two practice trials to familiarize them with the apparatus, test method, and confidence-rating procedure. They then had a study/test cycle on each of the three types of material. Before each exposure there was a verbal ready signal. For 30 subjects (group 1) each stimulus card was presented for 1 sec, while the remaining 30 subjects (group 2) received a .1-sec exposure. For all subjects there was a 5-sec interval between presentations. On test cycles, each test card was presented for 5 sec, with a 5-sec intercard interval during which each subject marked his answer sheet with a response and a confidence judgment.

RESULTS

Recognition of 'old' items

The proportion of old items recognized (hit rates) are summarized in Table 1. Entries are made for each combination of test-context condition, exposure duration, and class of materials. Each proportion was obtained by pooling over 30 subjects and three replications.

Table 1. Proportions of 'yes' responses to old target items in each test condition

Type of material	Test condition				
	Old context	Mixed context	New context	Permuted context	Removed context
Group 1 (1-sec exposure)					
Words	.86	.63	.60	.65	.79
CVCs	.78	.53	.67	.62	.69
CCCs	.68	.60	.49	.56	.61
Group 2 (.1-sec exposure)					
Words	.64	.58	.47	.66	.67
CVCs	.60	.47	.60	.56	.55
CCCs	.53	.48	.46	.48	.63

Visual inspection of the table shows that recognition for the sub-

jects in group 1 was best when they were tested with the original context and that any changes from the original display reduced recognition accuracy. Removal of contextual cues, however, appears to have been less detrimental to recognition accuracy than either the replacement of old components with new ones (mixed and new contexts) or the permutation of old context items. For both words and CCCs, recognition was least accurate in the new-context condition.

In contrast to that for group 1, the pattern of results for group 2 appears to have been different and more variable across materials. Changes from the old-context condition did not necessarily reduce recognition; in fact, for consonant trigrams, removal of context components may have enhanced recognition. Equally apparent is the fact that for word trigrams the removal or permutation of context components did not greatly alter performance as evidenced by the old-context baseline. However, as in group 1, the mixed- and new-context conditions indicated that word recognition may be disrupted when new components are presented in the test context. Finally, except for the permuted-context condition, the recognition of CCCs was extremely close to chance (.50). Apparently the recognition of CCCs by group 2 was greatly disrupted by the presence of any additional components in the test display.

An analysis of variance on the uncorrected raw scores for old items showed that the performance of group 1 was superior to that of group 2 [$F(1, 58) = 11.07, p < .01$]. Also, the nature of the materials proved to have a significant effect [$F(2, 116) = 8.55, p < .01$]. Similarly, there were significant differences among the test-context conditions [$F(4, 232) = 8.25, p < .01$]. Only the interaction of materials and context was significant [$F(8, 464) = 2.15, p < .05$]. A Newman-Keuls analysis was performed on the main effects of materials and context. This confirmed that the three materials (words, CVCs, CCCs) were significantly different from each other ($p < .05$), with words being the most accurately recognized and CCCs least. The old and removed contexts did not differ from each other, nor did the mixed, new, and permuted contexts. However, both old and removed contexts were significantly different ($p < .05$) from all three changed-context conditions.

Recognition of 'new' items

Table 2 summarizes the data on new items for each exposure group, each test condition, and for each class of materials. Although the data

were somewhat variable, the results suggest that of the three types of materials, the recognition of new *words* was least affected by the condition of testing. With CCCs and CVCs, however, the new-context condition definitely facilitated recognition of new items. There also appears to have been no systematic pattern of differences between the old- and removed-context conditions.

Table 2. Proportions of 'no' responses to new target items in each test condition

Type of material	Test condition			
	Old context	Mixed context	New context	Removed context
Group 1 (1-sec exposure)				
Words	.58	.68	.64	.63
CVCs	.59	.62	.71	.52
CCCs	.52	.56	.70	.64
Group 2 (.1-sec exposure)				
Words	.63	.63	.61	.56
CVCs	.54	.61	.79	.62
CCCs	.59	.57	.66	.62

An analysis of variance on the new responses showed no significant effect due to exposure duration nor any significant interaction containing exposure duration as one of its terms. However, there was a significant context effect [$F(3, 174) = 6.643, p < .01$]. There was also a significant interaction of materials and context [$F(6, 348) = 2.989, p < .01$]. A Newman-Keuls test on the main effect of context showed that the new-context condition was significantly better ($p < .05$) than any of the other three conditions, which did not differ among themselves.

The analysis of recognition in terms of signal-detection theory yielded a pattern of results very similar to that obtained with hit rates. Consequently, these data are not presented here.

DISCUSSION

The recognition of components of a compound display was clearly influenced by the stimulus conditions during tests. In general, the effects of contextual changes appeared more sharply with longer

stimulus exposures during training and with the more 'unitary' stimulus materials (words and CVCs). As always, a first set of data such as these is subject to multiple interpretations, a few of which are outlined below.

In many respects, the results appear to be consistent with traditional stimulus-response interpretations of contextual variation (Pan, 1926; McGeoch, 1942; Postman, Adams, and Phillips, 1955). In brief, this formulation explains the effects of contextual change in terms of three principles: (a) cue additivity, which suggests that the presence of multiple cues improves performance by providing a more effective eliciting stimulus, (b) cue reduction, which postulates that the removal of effective cues reduces performance, and (c) cue inhibition, which suggests that the introduction of new cues may disrupt performance due to the generation of strong competing responses. One possible point of discrepancy between our results and this view is that in group 1 the permuted-context condition produced lower recognition than the removed-context condition (cue reduction). Although the order of the contextual components was changed in the permuted-context condition, their presence should have provided some cue value. Finally, the expectations for recognition of 'new' items in various contexts are not explicit in this theory.

A recent and more complete development of association theory in terms of variable encoding responses has been proposed by Martin (1968). According to this view, any stimulus has the capacity to elicit a set of perceptual or encoding responses and each encoding then activates an association (response) if one exists. Now if, during presentation in the experiment just reported, the triplet components evoked encoding responses (which are essentially beyond the experimenters' control), it is likely that the manipulations of the test context changed those original encoding responses. Thus, the present method may have differentially reinstated the subjects' encoding responses as well as the nominal stimulus conditions.

Extrapolating from Martin's formulation, then, the alterations of context might have reduced recognition performance in at least two ways: (a) by preventing the arousal of encoding responses originally elicited by the contextual components and perhaps associated with the target component or its encoding responses and (b) by arousing novel encodings that were not recognizable as recent events and masked, changed, or interfered with the codes generated by the target item. Since these different encoding responses might be viewed as producing changes in context beyond the changes in the nominal stimulus con-

ditions, the largest decrement in recognition should occur for old items tested in the new-context conditions. Also, changing the order of the contextual units, as in the permuted-context condition, might be expected to alter the subjects' encoding responses to the target item and reduce recognition. Similarly, the recognition of new items embedded in various context conditions might be influenced by the relative familiarity of encoding responses to context items.

In general, both theories outlined above seem to offer reasonable interpretations of the effects of altered stimulating conditions on recognition. However, a more complete explanation of contextual effects may require explicit representations of what the subjects extract and store, as well as of how the retrieval process contacts 'traces' deposited by prior experience. The present results suggest that under the test conditions used here, the recognition response is based upon some weighted combination of (a) the total pattern of stimulation present, (b) the particular features (acoustic, semantic, and so on) of the target component being tested, and (c) the characteristics of the untested elements in the display.

Depending on the type of learning materials involved, it might be assumed that subjects extract and store, in a certain proportion of cases, representations of the whole pattern or triplet (whole-pattern representations) as well as representations of the individual units or parts of units (component representations), given that a subject has had enough time for multiple sampling from the stimulus display. The multiple representations may enter separate 'stores' or 'dictionaries,' which can be searched independently of one another. Thus, depending on how much encoding or 'feature extracting' was possible, and therefore, on how much was stored at each 'level of integration,' constraints may be placed on the type of retrieval processes possible.

From this 'multiple dictionary' hypothesis, one might surmise that when the subjects were tested under the removed-context condition in the present experiment, they were forced to retrieve on the basis of component representatives—to do so whether or not they had a representative of the whole pattern stored. Also, during their retrieval attempts, there were no context items present to influence their search process or to 'weight' or 'bias' their final recognition response. Consequently, even with long sampling times, one might expect an intermediate level of performance for component recognition in the removed-context condition.

At the other extreme, the original pattern of stimulation was presented on tests for items in the old-context condition. In this case, the

subjects could attempt retrieval in at least one of the two modes, depending on the degree of 'multiple dictionary' representation in storage. (For the present we will not speculate as to whether alternative retrieval modes might occur serially or in parallel.) First, the subjects might attempt to match the test stimulus by searching on the basis of the whole-pattern representation, since from a successful pattern match, they could infer that the component unit was 'old.' Failing this, they could attempt retrieval on the basis of individual components or subunits and find a match in this mode. Also, if they used the component mode in the old-context condition, there were no irrelevant context components in the test display to interfere with searching or to negatively bias their final recognition decision. Thus, one might expect the best recognition performance in the old-context condition, insofar as both pattern and component representations were stored. This would be expected with longer exposure durations and with more meaningful stimulus units.

In contrast, correct recognition responses in the case of items tested in the mixed-, new-, and permuted-context conditions could not be successfully produced on the basis of whole-pattern recognition. Since the test display was a triplet, the subjects might sometimes attempt to search in this mode, but these attempts should simply waste retrieval time. Thus, under these conditions of testing, the subjects must ultimately retrieve on the basis of component representations. Since the 'irrelevant' context components present might disrupt both search and decision processes, one would expect recognition performance to be lowest under these conditions.

Although the 'multiple-dictionary registration' hypothesis presented above is rather incomplete, it may serve as an initial theoretical description of the subjects' behavior in the present experiment.

Notes

The authors express their gratitude to W. Gilson for assisting in the collection of the data. Received for publication June 2, 1971; revision, August 11, 1971.

References

- Bower, G. H., Clark, M. C., Lesgold, A. M., and Winzenz, D. 1969. Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior* 8:323-343.

- Light, L. L., and Carter-Sobell, L. 1970. Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior* 9:1-11.
- Martin, E. 1968. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review* 75:421-441.
- McGeoch, J. A., 1942. *The psychology of human learning*. New York: Longmans.
- Pan, S. 1926. The influence of context upon learning and recall. *Journal of Experimental Psychology* 9:468-491.
- Peterson, L. R., and Peterson, M. J. 1957. The role of context stimuli in verbal learning. *Journal of Experimental Psychology* 53:102-105.
- Postman, L., Adams, P. A., and Phillips, L. W. 1955. Studies in incidental learning: II. The effects of association value and of the method of testing. *Journal of Experimental Psychology* 49:1-10.
- Tulving, E., and Pearlstone, Z. 1966. Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior* 5:381-391.
- Tulving, E., and Thomson, D. M. 1971. Retrieval processes in recognition memory: Effects of associative context. *Journal of Experimental Psychology* 87:116-124.
- Underwood, B. J., and Schulz, R. W. 1960. *Meaningfulness and verbal learning*. Philadelphia: Lippincott.
- Weiss, W., and Margoulis, G. 1954. The effect of context stimuli on learning and retention. *Journal of Experimental Psychology* 48:318-322.

BOOK REVIEWS

Julian Hochberg
Columbia University

Kinesics and Context: Essays on Body Motion Communication

By Ray L. Birdwhistell. Philadelphia: University of Pennsylvania Press, 1970. Pp. xiv, 338. Paperback, \$3.95.

Ray Birdwhistell is an anthropologist who for more than 20 years has taken as his focus of interest the phenomena of human communication. His concern is with the patterns of human behavior that people manifest when they communicate, and his aim is to describe these patterns. Birdwhistell, who recognizes that communication is mediated by a complex system of behavior that includes speech, movement, and the relations of the communicators to the physical environment, is to be regarded as one of the earliest exponents of what has been distinguished (Duncan, 1969) as the 'structural' approach to communicational behavior, and as such, he has had an important influence upon a number of workers in this field.¹ His main focus, however, has been on the way that body motion functions in communication. In his attempts to set this topic on a systematic footing, he is a pioneer, and for many years he was the only scholar engaged in the task. The book reviewed here brings together Birdwhistell's more important writings. Hitherto, these were hard to obtain, so the appearance of this volume is very much to be welcomed.

The newcomer to Birdwhistell should be forewarned. Anyone who opens this book expecting to find in it a coherent exposition of the role of body motion in communication will be disappointed. If he is looking for philosophical and methodological guidance, however, he may derive much. Birdwhistell is not a systematic writer. In these essays he shows us something of what a systematic study of communication behavior might look like, and he provides the foundations for an approach to the subject that will be of great value if properly understood and applied; in places throughout the book we get brilliant observations on communicational behavior. We do not get a well-worked map of the territory. Despite some reviewers, however, I see no reason why we should have expected such a map. It is only within the last five years that anyone other than Birdwhistell himself, besides a mere handful of workers, has even begun to take this field seriously.

The book is not easy to read. Though Birdwhistell at times writes with simplicity and force, he is often dense, technical, and abstract. To understand what he is saying, the reader must often work rather hard, especially if he is unfamiliar with the approach. In this review I shall try to summarize what I take to be the most important features of Birdwhistell's approach, and I

shall contrast his approach to that commonly taken by psychologists who have interested themselves in body motion. I shall then deal with Birdwhistell's more specific statements on kinesics. In doing this I hope not only to provide a guide for those who would read this book but also to outline an approach to the study of human communication, an approach which is only now beginning to take shape but one which will, when applied in earnest, prove to be of considerable importance.

Birdwhistell is interested in the phenomena of face-to-face communication in people. He is interested in the order or pattern that can be observed in the ways in which people relate their behavior to one another when they are in each other's presence. For him this order or pattern is communication: "when we talk about communication, . . . we discuss it as a complex and sustaining system through which various members of the society interrelate with more or less efficiency and facility" (p. 12). The aim of a science of human communication, in his view, would be the systematic description of the systems of behavior by which this interrelatedness is brought about. For Birdwhistell, thus, communication is not something that we may or may not achieve; it is not something that we may or may not do of our own choosing, for whenever we are in the presence of another, to the extent that there is an interrelation between our behavior and that of the other, to this extent communication is going on. For Birdwhistell, communication is to be viewed as a system with a structure that can be described independently of the behavior of particular participants. Particular participants, on given occasions of interaction, might be said to make use of this system. The aim of Birdwhistell's endeavor is to give us an account of this system.

This systemic view of communication is not original with Birdwhistell. It received an earlier formulation in the work of Gregory Bateson, to whom Birdwhistell acknowledges his debt. It is implicit in the views of such anthropological linguists as Sapir and Kenneth Pike and others.² Birdwhistell gives it a forceful expression in this book, however, and he shows clearly what its implications are for a methodology for the study of communicative behavior.

By this systemic view of communication, anything that anyone does in the presence of another must be considered as potentially part of the system. Speech and gesture, posture and orientation, touch and relative position in space—all must be taken into account if we are to comprehend communication. We cannot at the outset of our investigations decide not to attend to certain aspects of behavior. So long as it is detectable by the other, it must be presumed communicative until proven otherwise. That is, though we must exclude nothing at the outset, one of the outcomes of our work would be to show what aspects of behavior are *not* part of the system. As Birdwhistell says of body motion, "all of the observed shifts of the human body are not of equal significance to the human communicational system. . . . Which particular behaviors are of patterned communicative value . . . can be determined only by the systematic investigation of the behavior in the communicational context. . . . Our problem is to describe the structure of body motion communication behavior in a way which allows us to measure the significance of particular motions or complexes of motions to the communicational processes" (pp. 76-77).

It will also be seen that we cannot regard any one aspect of behavior as necessarily 'more communicative' than another. For instance, when two people talk together, the physical presences of each to the other, their respective orientations and their postures, are all part of the conditions that make the talk possible. They must be considered an integral part of the communicative system, along with the words that are uttered and the gestures that are performed. From Birdwhistell's point of view, those aspects of behavior that establish and maintain relatedness are just as important as those behaviors that transmit new information from one person to another. To focus on language alone or on body motion alone is thus to focus on only a part of the system, and until its relationship with the other parts is also understood, our understanding of communication will be limited indeed. It makes no sense to speak of 'verbal communication' and 'nonverbal communication.' There is only communication, a system of behavior patterns by which people are related to one another. This does not prevent us from looking at the different aspects separately, and indeed it proves useful to do so, up to a point. Birdwhistell's original and still abiding interest is in the part played by body motion in communication, yet he is fully aware that it is but a part and that the decision to study it separately is but a matter of investigatory convenience.

The focus, then, is on what behaviors people characteristically engage in when they interact. Since we do not know what these behaviors are, we must look and see. Most often, an investigator with this orientation will seek to gather records on film or video tape of occasions when people are present to one another and then, by patient and detailed watching, he will try to describe the elements of behavior that occur and the way these elements are patterned in relation to one another. Gradually, as he accumulates examples, he will be able to state the contexts within which the elements he has isolated may be found, and from such statements specify the ways in which these elements function in the communication system he is studying.

The most rigorously developed methodology consonant with this approach is that of descriptive linguistics. Birdwhistell recognized this at an early stage in his career. He sees the methods of descriptive linguistics as a model for those who would examine the structure of other systems of behavior in communication. In developing kinesics he has adapted much from these methods, and he has also adapted much from its terminology. He acknowledges particular debts to the work of H. L. Smith, George Trager, and Norman McQuown in this regard.

I will return later to a more detailed discussion of Birdwhistell's efforts to demonstrate how much of body motion, as it functions in communication, has a structure analogous to that of spoken language. For the moment I would like to point out that suggesting a 'linguistic approach' to the analysis of systems of communicative behavior other than language does not mean supposing that these other systems are languages. It means, rather, adopting the *level of analysis* at which linguists operate when they approach speech and seeking for their *mode of explanation*. Such an approach does assume, of course, that when people interact they make use of a repertoire of behavioral forms that they share with others and use in accordance with sets of shared rules. This does not presuppose the nature of these forms,

however, nor does it presuppose the nature of their rules of patterning. These may be quite different from those that may be found in language.

In recent years many psychologists have become interested in what is often rather unfortunately referred to as 'nonverbal communication,' and they are thus looking at many of the same phenomena that interest Birdwhistell. However, almost all who are working in this area differ from Birdwhistell in that they use an experimental approach. They also ask a different kind of question. It may help to show in more detail what the nature of this difference is. I will do this by contrasting the way 'facial expression' has been traditionally investigated in psychology with the approach to the same phenomena that Birdwhistell would follow.³

In the psychological approach to the study of the face, there are two principal methods that have been followed.⁴ On the one hand, subjects have been placed in situations in which they have been presented with various stimuli presumed to arouse different emotions, and the behavior of their faces has been recorded in the hope that the patterns appearing there would differentiate in a reliable way the different stimuli presented. In the other, more common approach, the question has been whether people can reliably distinguish facial expressions in terms of different emotions the faces are presumed to be expressing. The usual method here has been to present subjects with photographs of faces in various poses and to ask them to label the photographs in terms of the emotions supposedly being expressed. A number of other methods have been used from time to time, but in all of this work, it will be noticed, it is assumed that the various poses that we can observe on the faces are *expressions of emotion*. The ultimate interest of almost all workers in what is by now a very long tradition has been in the emotions. Interest in the face seems to be limited to the extent to which it can indicate the inner affective states of the person whose face it is. In the experiments in which people are asked to make judgments of photographs of faces, accordingly, the subjects are always directed to consider *only* that the faces are expressing emotion.

In Birdwhistell's approach, one begins not with an interest in the emotions but with an interest in the face itself. One would ask first, What are the various things the face does? One would then proceed to determine, by careful observation, the various settings in which repeatedly observed units of facial behavior can be seen to occur, and the question would always be, What functions *for the interaction* do these differentiable units of facial behavior have? From this point of view, in other words, the question of what inner state is supposedly made manifest in the face is not relevant. What is relevant is what difference different facial displays make to the organization of the occasions of interaction in which they occur.

Remarkably little research on facial behavior from this perspective has been carried out. There is some research on the infant smile, summarized by Bowlby (1969), in which the role of the smile in promoting the mother/child interaction has been examined. Van Hoof (1962, 1972), Blurton-Jones (1967), and Grant (1968, 1969) have reported observational studies of facial (and other) behavior in which the principal interest is in its signaling function. Most of this work has been on children of nursery-school age. Rosenfeld (1966) and Reece and Whitman (1962), working within an

experimental approach, have reported studies of the influence of smiling on other aspects of interaction. Birdwhistell himself, in the book here reviewed, offers an essay on the smile. In it he points out how the smile in itself is a highly complex phenomenon, with a wide variety of social functions, and that to consider it merely, in his phrase, as a "visible transform of an inner physiological state" would be to miss entirely its significance as a social signal. He also refers, in various places in the book, to how movements in the face may be brought into play in association with speech, and to how these movements, along with movements of the head and limbs, have a complex relationship both with the structuring of the speech as an activity and also with its content. Facial displays can serve to mark out points of emphasis in speech, they can serve to mark off whole segments of speech as distinct units or as contained or embedded units, and they can also provide a sort of commentary on what is being said. If we watch the faces of listeners, too, we can see that nods, smiles, frowns, raised eyebrows, appear frequently in some circumstances and are an important part of the repertoire of the listeners' behaviors and serve to regulate the behavior of the speaker. Though no systematic data are offered by Birdwhistell, his occasional observations are enough to make it clear that work along these lines will be highly illuminating.

Birdwhistell, concerned as he is with the interactional use of the face, and as an anthropologist, aware of how communicative behavior is patterned differently in different societies, has expressed the view that there are no facial expressions that have the same social meaning all over the world. Specifically, he says also that there are no universal gestures of emotional expression. Ekman (1971) and Eibl-Eibesfeldt (1970), though from very different perspectives, have recently challenged this view. Ekman (1971) has gone to considerable lengths to demonstrate that different emotions can be reliably recognized from still photographs of differently posed faces in the same way in a number of widely differing societies. On the basis of these and certain other studies, he claims that there are distinctive patterns of muscular activity in the face that are characteristic of different emotions and that these patterns are universal, part of the expressive equipment we are all born with. Ekman claims that his findings on facial expression constitute a refutation of Birdwhistell, who, he says, "cannot admit the possibility of universals in facial expression and maintain his major central claim that facial and body behavior is a language" (Ekman, 1971, pp. 2-3).

In reply to this, there are two points to be made. In the first place, it should be clear that even if there are "universal forms of emotional expression," this would not mean having to give up the idea that "facial and body behavior is a language." As I have tried to make clear, Birdwhistell's use of the linguistic level of analysis and his adaptation of linguistic terminology reflect his view that it is appropriate to study the behaviors that are used in communication as a system in its own right. There is no *logical* reason we cannot adhere to this level of analysis and recognize that some features of this system are not specific to any given culture.

Secondly, it should be pointed out that Ekman's investigations follow the traditional approach—that he is interested in the face only insofar as it seems to allow one to apprehend the inner states of the individual. He does

not ask, as Birdwhistell would, How is the face used in interaction? Nor do his findings contribute to any answer to this. When Ekman uses the phrase 'expression of emotion,' he uses it in the traditional sense that the emotion is something 'inside' the individual that 'comes out' on the face in a particular way. Birdwhistell, on the other hand, approaching behavior in terms of its communicative function, rarely discusses emotion as such. He provides no explicit analysis of the concept, but from his perspective it would seem that his concept of emotion would be quite different from Ekman's. Birdwhistell would describe emotion not as an inner state but as a form of behavioral relatedness between individuals. Those displays that we have traditionally called displays of emotion he would describe not in terms of what they express but in terms of how they function in interactions. That is, affect displays would be seen as displays that bring about certain kinds of adjustments in the behavior of others, adjustments distinct from adjustments that occur in association with other kinds of signals. From this viewpoint, it may well turn out that the forms of facial display that Ekman (and also Eibl-Eibesfeldt) find to be universal are nonetheless embedded in interactional sequences in characteristically different ways in different cultures, so that something like Birdwhistell's original statement may still turn out to be correct. As it is, Ekman himself recognizes that societies differ in the kinds of rules they have to govern the display of feeling, rules which include rules about what kinds of facial display a person may use in different situations. A final resolution of this difference between these two investigators will come only when we have a much better understanding of the signaling function of the face, on the one hand, and, on the other, a much better understanding of what 'emotion' is, not from the standpoint of its physiology, but from the standpoint of its place in interaction. In all probability, it will turn out that there is much less difference between Ekman and Birdwhistell than there appears to be. At the moment, it seems likely that they differ mainly because they are not really talking about the same things.

Birdwhistell is very insistent that we can never state the meaning of a behavioral unit except in terms of its place within a context. It may be true that whenever we see a certain facial expression, no matter what the context, we can be sure that the person who shows it has a certain inner state. In this case, there is a sense in which we can say that 'facial display X means inner state Y.' But this is only one kind of meaning, and it is not the kind of meaning with which Birdwhistell is concerned. For him, to ask what a given unit of body motion 'means' is to ask what its *use* is. Questions such as, What does it mean when a person smiles? or What does it mean when a person raises his fist and shakes it back and forth? can only be answered by giving a list of the environments in which these forms are generally seen. In other words, we can answer these questions in terms of the range of use.⁵

Birdwhistell says that his realization that even 'well-known' gestures like the army salute depend on their context for their meaning in any instance was an important step toward his view that gesture and other aspects of body motion could profitably be approached by linguistic methods. He came to the view that some aspects of body motion, at least, could be seen to be organized in a way that is analogous to language. This view is perhaps best expounded through an example.

If we see a man lift his clenched fist and move it rapidly forward and back two or three times, we cannot conclude, as we might be tempted to, that he is 'angry.' He may be engaged in a friendly salutation with someone. He may be engaging in a pictorial gesture, as he describes someone he knows as being tightfisted. His raised and shaken fist may be his reply to a question about what a mutual acquaintance did in a certain situation (with this gesture he makes a statement which in this context means something like 'he was very tough'). The shaken fist as a 'gesture of anger' is only one among many of its uses; and even as a 'gesture of anger,' it may not mean 'I am angry' but rather 'he was angry,' or 'I would be angry if . . .,' or 'I was angry then,' or even 'I mean anger, not fear.'

Now, in all of these cases, not only is the general context of the appearance of this gesture different, but also the gesture itself, the raised and shaken fist, is only one element within a configuration of several elements, a configuration which is different in each case. When the raised fist appears in the salutation, for example, it may be accompanied by an openmouthed upper-toothed smile and raised eyebrows (a facial display common in greeting situations). When the raised fist appears to illustrate the idea of tightfistedness, it may be accompanied by the statement 'he was very tightfisted,' the speaker's brows fully raised and his lower and upper eyelids drawn together to produce 'slit eyes.' When it appears as a gesture meaning 'I am angry,' it may occur with the head tilted forward and with the brows drawn together over the bridge of the nose, the mouth closed and the lips pressed tightly together and slightly rolled inward. Each of these cases differs in other ways, too, such as in the position of the fist in relation to the head: in the 'angry' case, it was held rather close to the face of the fist shaker; in the salutation example, it was held above the head, somewhat to the right, while the body of the person was facing at right angles to the person he was saluting.

These examples show how a complex statement made in body motion alone or in body motion and accompanying utterances can be made up of a number of separable elements. The raised clenched fist is one such element. The various facial displays that can accompany it are other such elements. These facial displays can also occur in other contexts, independently of fist shaking, and indeed consideration of these facial elements will show that they themselves are combinations of elements of facial behavior that can occur separately. In describing the behavior of the face, it is convenient to consider at least the brows, the muscles around the eyes, the eyelids, the nostrils, and the mouth as separable 'areas of articulation,' each one of which has its own repertoire. Facial configurations then arise from different combinations of elements in these repertoires. It is worth noting that though Birdwhistell was the first investigator to try to deal with the face in this way,⁶ similar but quite independent attempts were made by van Hoof (1962) for primate faces and by Ekman and Kendon (1964). More recently, Grant (1969) has developed descriptions of different faces along similar lines, and Ekman's visual 'dictionary' of facial elements observes some of the same principles (Ekman, Friesen, and Tomkins, 1971).

The basic idea I have tried to illustrate in the discussion above is that different parts of the body can be viewed as having repertoires of elements and that these elements can occur together in different combinations to

produce more complex configurations. These more complex configurations, such as the gestural statements described above, can themselves be seen to behave as units and to occur in complex relationships with each other. It will be further seen that these more complex units are truly 'emergent' units, in that the alteration of only one element may completely alter the whole complex.

The idea that we can talk of elements of body motion that occur in different combinations, combinations which themselves behave like elements in *other* combinations, is one of the most central ideas in Birdwhistell's conception of kinesics. The parallel with conceptions developed in structural linguistics will not be missed, and Birdwhistell has tried hard to develop this parallel in some detail. Thus he suggests that the smallest elements be referred to as kinemes, named in direct analogy to phonemes. A kineme, for Birdwhistell, is a class of movements whose members may vary among themselves, but this variation is not such as to make a difference to the larger units of which it is a part. Thus, in any two examples of a facial display that includes raised brows, let us say, so long as the facial display is 'the same' in function despite any actual difference in the degree to which the brows are raised in the two cases, then, in the terminology Birdwhistell has developed, the two cases of brow raising may be regarded as *allokines* of one another and raised brows in general as a *kineme*. Other examples of kinemes here would be knitted brows and puckered brows. Birdwhistell provides a list of many of the kinemes he has established, though unfortunately he does not give any detailed data about them.

Kinemes, in Birdwhistell's conception, combine in various ways into configurations that he calls *kinemorphs*. The various facial displays I described above in the examples of the raised fist are examples of the sort of thing Birdwhistell means by kinemorphs. The total configuration, fist plus facial display, which I called a 'gestural statement,' is a combination of kinemorphs, which Birdwhistell would call a *kinemorphic construction*.

Such gestural statements or kinemorphic constructions as 'fist shaking in salutation' or 'fist shaking to illustrate tightfistedness' occur, of course, in larger contexts, and to handle these Birdwhistell suggests some additional ideas. If we think of the 'fist shaking to illustrate tightfistedness,' for instance, we note that this construction occurs in association with a linguistic construction that, in a very loose sense, it may be said to illustrate. Prior to this, and following it, other linguistic and kinesic constructions may occur, all, for example, while the speaker is sitting in a particular posture in an armchair that is at an angle of roughly 90 degrees to another armchair in which another person is sitting. In other words, this linguistic-kinesic construction occurs within a conversation. When people 'have conversations,' they arrange themselves in space in relation to one another in one of a limited number of ways, and the sustained postural and orientational arrangement here functions as a frame within which the succeeding constructions are seen to occur. This frame serves to bind together such sequences of constructions. It serves to cross-reference them, as Birdwhistell would put it, as constructions belonging to this conversational unit. For the individual, Birdwhistell uses the term *stance* to refer to those features of behavior that have this cross-referencing, binding, or framing function. Some of the

phenomena at this level of functioning in the kinesic stream have been explicated more fully by Schefflen (1964).

A stance, it will be seen, is not a construction of smaller elements but an element in its own right, an element which, as I have indicated, functions to 'frame' sequences of constructions. It is important that this be brought out, for some readers of Birdwhistell get the impression that he regards all of body motion, as it functions in communication, as being built up of elementary units, or kinemes. On the contrary, the constructions that do emerge from combinations of kinemes must be joined together and they must always occur within a 'transfixing' frame, which is supplied by such enduring aspects of behavior as posture and relative position in space.

Birdwhistell also recognizes some other aspects of body motion that cannot be handled in terms of constructions. For example, he points out that although we may describe a piece of an individual's behavior in interaction in terms of a succession of constructions (together with the frames within which these construction sequences occur), there is also considerable variation in how these units are enacted. We all recognize this in an informal way. When we recognize a 'weak handshake' or a 'vague smile,' for instance, we recognize the occurrence of a construction for what it is, but we also remark on the quality of its performance. Birdwhistell thus distinguishes *motion qualifiers*, which refer to the intensity, the duration, and the range of movement in a construction. These qualifiers may inform the whole of a person's performance (and thus may signal something that we call 'mood'). They may also inform different constructions in different ways and in doing so they may thus modify the meanings of given constructions. For instance, X may shake hands 'vigorously' with Y, but 'limply' or 'briefly' with Z. In doing this he may be said to 'state' differences in his attitudes or feelings about Y and Z.

In the exposition above, I have tried to deal with what seem to me to be the most important features of Birdwhistell's conception of how body motion is organized in interaction. Most of Birdwhistell's ideas on this particular aspect of the phenomena of communication are to be found in chapter 20 of the collection under review. This chapter represents his most comprehensive attempt to state, in a systematic way, his views on the organization of body motion, and it well repays careful study. Unfortunately, however, it provides very few examples, and for some of the most important ideas he gives no examples at all. For this reason, the chapter is very difficult to read. It is of the greatest value—as is indeed the whole book—to those who have already worked in kinesics. However, this chapter opens with a superb analysis of a specific interactional event in which two young men try to 'thumb a ride' from a passing car. More than anything else in this book this example should give the reader a clear feel for the sort of analysis Birdwhistell seeks to advocate. Birdwhistell would have done his readers much greater service, however, if he had provided many more examples of this type, using them to show just how the concepts he develops in the abstract are to be applied.

The phenomena Birdwhistell exemplifies in greatest detail have to do with how movement is patterned in relation to speech. He apparently sees this as a somewhat separate phenomenon, and it is something he has only looked

at closely in recent years. Part Three of the present collection includes three papers on this topic.

For example, when someone speaks, he always moves more than just his mouth. We can see eyeblinks, eye movements, brow movements, and movements of the mouth that are not part of the articulatory movements required by what is being said. We may often see head movements, and not uncommonly, other parts of the body, such as the hands and arms, are moved as well. These movements are highly patterned, and a close analysis shows that they are so highly interlinked with speech that they must be considered as part of a total act which includes speaking.

The way in which these movements relate to speech is a very complex matter, and no simple statement can be made. It may be shown, for instance, that syllabic units, phrasal units, and units of speech comprised of groupings of phrases are all marked off in distinct ways by movement units. This aspect of the phenomenon referred to by Birdwhistell has been demonstrated in detail in some recent work of my own (Kendon, 1972). However, movements that co-occur with speech also relate to the content of what is being said in quite complicated ways. People may parallel their speech output with gesticulations to illustrate what they are saying, either by drawing pictures in the air of what they are talking about or, if they are talking about something abstract, by making movements that appear to make partially concrete their abstract talk (as when someone, in making a succession of points in an argument, places his hands in a succession of different places in front of him as he talks, a new place for each point). Some of these aspects of the relations between gesticulation and speech are treated by Efron (1942). They have been summarized recently by Ekman and Friesen (1969).

In the present collection, Birdwhistell does not deal in detail with either of these two examples of ways in which movement may be related to speech (although he is, of course, aware of them) but rather with two aspects of how movement may mark some of the internal details of the utterance. These two aspects are linguistic stress, which he shows is also often marked kinesically, and the kinesic marking of certain syntactic classes.

In his treatment of linguistic stress, Birdwhistell shows, by a series of examples, how contrastive movements can often be observed to co-occur with the points of linguistic stress in an utterance. Such movements may be very slight and consist in small movements of the brows, eyeblinks, or slight movements of the head. It is important to note that Birdwhistell does not mean, as some (Dittman and Llewellyn, 1969, for example) have thought, that there will be 'more movement' at points of stress than elsewhere. There is movement all the time. What Birdwhistell means is that there will be a contrasting pattern of movement at points of stress. If the reader attunes himself to this phenomenon by working through Birdwhistell's examples—by repeating the utterances in front of a mirror, or by getting a friend to say them—he can easily confirm for himself much of what Birdwhistell is saying about kinesic stress. Systematic analysis, of course, which will be the only way Birdwhistell's finding can be further investigated, requires the use of sound-synchronized film and suitable film-analysis equipment.

The contrast in movement, or the discrete movement unit that co-occurs

with points of stress in an utterance, is not merely contrastive, however. Birdwhistell shows that there may be a systematic relation between the direction of movement and the kind of word receiving linguistic stress. Thus in utterances in which the pronouns 'I' or 'me' or such deictic words as 'here' or 'this' are stressed, the movement made as they are uttered is a movement toward the self. With similarly stressed words such as 'you' or 'him' or with words such as 'there' or 'that,' the movement is away from the self. Words that have a temporal reference take a forward movement if the reference is to the present or future, but they take a backward movement if they refer to the past. Birdwhistell also says that movement contrast can be shown to mark plurality, the tense of verbal phrases, and also prepositions. Once again, if the reader works with the examples Birdwhistell gives, and so attunes himself to the phenomenon, he can pick up some of it as he watches others speak. Systematic analysis using sound film and analysis equipment is much needed, however. Examples of some of these kinesic marker phenomena are described in some detail by Kendon (1972).

The movements that co-occur with speech often have multiple functions. In all the examples Birdwhistell gives—examples of movements that function both as markers of stress and as markers of types of phrase—the movements appear as a kind of parallel to the speech output. Often, however, such movements actually add to what is being said. One example from my own collection must suffice here to illustrate this. A husband, at the end of the day, was sitting in the living room talking with his wife in a casual way about what the children had done during the day. At one point he said: "They made a cake, didn't they?" (He was referring to something he had been told by one of the children earlier, that they had made a cake of mud in the garden.) As he said 'cake,' this received not only primary linguistic stress but also kinesic stress, in a rightward tilt of the head. The tilt of his head was toward the window overlooking the garden. The direction of the head tilt thus built into the utterance a reference to where the cake had been made and hence to the *mud* cake the children had made, not some other cake. Part of what was conveyed by his utterance was thus conveyed not in words or tone of voice but in the directionality of the kinesic stress marker that co-occurred with his words. By the way a speaker patterns his movements with his speech, therefore, he may 'pack in' far more content than is apparent if we consider the words alone. Examples of this sort are not hard to find if one makes one's observations with care, and they provide a rather convincing demonstration that we must consider movement as well as speech if we are to understand what is entailed in what we somewhat loosely refer to as an act of speech.

Throughout this review I have used the term body motion, as Birdwhistell himself often does, as if it can be regarded as a unified system. In various places in these essays, however, we can find remarks that suggest that this may not be Birdwhistell's view. I have already said that Birdwhistell appears to regard the patterning of movement in relation to speech as a phenomenon that can be given separate treatment. In several places, he makes reference to systems of sign language and to such specialized systems of gesture as may be found in Indian classical dancing, but only to show he is not concerned with these. He also, in one place, seems to

shunt to one side those elaborate demonstrative or pictorial movements that people use when they describe something: a man describing how he played and landed a trout, for example. Birdwhistell says, "At the moment I am inclined to regard such behavior as examples of derived communicational systems. As such, they are not the primary subject matter of kinesics at present" (p. 126). Birdwhistell thus implies that kinesics deals with only a part of communicatively significant body-motion phenomena. However, he never makes clear, in so many words, what this is. Presumably this would become apparent as the study of the whole system of communicative behavior progresses.

It seems likely that ultimately we will have to think of the system of human communicative behavior as analyzable into a number of subsystems, subsystems distinguished in terms of which aspect of an occasion of interaction they 'take care of.' For a conversation, for example, one aspect would be the conversation as an occasion. By a variety of behavioral devices, each participant will signal to the other that he is a member of the conversational unit. Another aspect of the system will have to do with the situational role each participant has in the conversation, and another set of behavioral devices will refer to this. Yet another aspect of the system will have to do with ideas or thoughts or stories that the participants are conveying to one another. Each of these interactive systems will be seen as making use of one or more of a number of available codes, of which language is the most obvious example.

There are also a number of body-motion codes, however. The patterning of movement in relation to speech that Birdwhistell describes may be one of them. The postural system that Scheffen (1964, 1965) has described might be regarded as another. The system of kinemorphs and kinemorphic constructions that Birdwhistell refers us to may be yet another. Specialized sign languages constitute additional ones. Each of these codes has emerged in response to a somewhat different demand made by the interactional situation, and it presumably will be found to have properties tailored to match. The principles by which one of these is organized will not necessarily be generalizable to another. These codes, it should be noted, are probably best thought of as separable only in the analytic sense. In practice they are interrelated in highly complex ways, which are far from being understood.

Such is the view of the organization of communicative behavior implied by Birdwhistell. It seems to me, at least, that the picture which finally emerges will look something like this. But whatever the details of that picture turn out to be, there is little doubt that it will owe much to the perspective Birdwhistell provides, and for this reason, if for no other, his writings deserve careful study.

To conclude, I must return to the book itself. Unfortunately, taken as a whole, it is not very satisfactory. This is because the editor, Barton Jones, has tried to make a coherent whole from an assemblage of very diverse pieces. He has done this by arranging the papers in topical groupings and omitting passages that are repetitive. Sometimes he has forged a new chapter by putting together different papers. The result, however, is not a book, but a collection of fragments. These fragments, furthermore, are not presented

to us in chronological order, even within the topical sections. So it is impossible to follow the development of Birdwhistell's thinking. Yet, if we compare earlier papers with the later ones, it is clear that Birdwhistell's thinking has evolved considerably, particularly in regard to his view of the nature of communication.

If the reader tries to read the book consecutively, therefore, he will come across a number of inconsistencies. Unfortunately, these are nowhere explained. For example, Birdwhistell's use of the term *kine* undergoes a change. The use he makes of this term in chapter 26 (written about 15 years ago) is different from the use he makes of it in papers written later— notwithstanding the fact that these papers appear earlier in the book. Some editorial apparatus to iron out these problems is badly needed. There are also many inconsistencies between the bibliography in the back (in itself a very useful one) and references in the text. There is no index, which is especially unfortunate in this book, for in order to trace Birdwhistell's views on a given topic it is necessary to jump around in the text. In view of the importance that I believe Birdwhistell's views have, these drawbacks are unfortunate, especially because he is often far from easy to understand. Nonetheless, had it not been for the insistence of Erving Goffman (the coeditor of the series in which this book is published) and the work done by Barton Jones, we might never have had these papers in print. For this we must be very thankful.

Adam Kendon, *Bronx State Hospital, New York*

Notes

1. Others who share this perspective and who have been influenced by Birdwhistell include Erving Goffman (1963, 1971), Albert Scheffen (1964, 1965), and E. T. Hall (1959, 1966).

2. See Bateson and Mead (1942) and Reusch and Bateson (1951). Bateson, with Margaret Mead, also pioneered the use of film for the recording and analysis of human social behavior, which was to become a central technique in Birdwhistell's methodology. For the work of Sapir, see Mandelbaum (1963); for that of Pike, see Pike (1967).

3. I do not mean to imply that all psychologists who have worked on 'nonverbal communication' have been concerned with the emotions. However, as Duncan (1969) has pointed out, in most of this work the investigator looks for how the particular behavioral parameters he is interested in correlate with, and so can be regarded as indexes of, other aspects of the psychological state of the subjects. For example, most of the work on direction of gaze in interaction has been of this type: typically, measurements are made of how much a person looks at an interlocutor and these are then related to such things as amount of liking or degrees of embarrassment. Alternatively, some aspect of the behavior of one interactor is systematically varied, and an effort is made to find out how someone else's judgment of the individual is affected. These are questions of a different sort from those asked in Birdwhistell's perspective.

4. For a recent comprehensive review, see Vine (1970).

5. In Birdwhistell's words, " 'What does X mean' translates as 'How do you know the place of this phenomenon in that larger pattern you are describing?' It is this question that the methodology of kinesics is designed to answer" (p. 185). This view that the meaning of an item is accounted for when we have described its context of occurrence is widely held in linguistics. Compare, for example, W. Haas (1956, quoted in Waldron, 1967, p. 207): "To... what is the meaning of (say) *house*? we shall answer: It is its distinctive occurrence in certain linguistic frames, or (as I would prefer to say) in certain sentential functions such as (using x, y, z, etc., to mark positions) This x is big, I wish to buy a y, John's z, etc. *House* may be substituted in x, y, z, etc.; that is its meaning." In some ways, Birdwhistell's view of meaning is closest to that known as the context-of-situation theory, originally stated by Malinowski (1923) and later developed by J. R. Firth (1957).

6. Birdwhistell has dealt with all areas of the body, of course, not just the face. Two of his notational systems are given as appendixes to the volume here under review.

7. See also the work of William Condon (Condon and Ogston 1966, 1967).

References

- Bateson, G., and Mead, Margaret. 1942. *Balinese character, a photographic analysis*. Special Publications of the New York Academy of Sciences, vol. 2, ed. W. G. Valentine. New York: New York Academy of Sciences.
- Bowlby, J. 1969. *Attachment and loss*. Vol. 1, *Attachment*. London: Tavistock Institute; New York: Basic Books.
- Condon, W. S., and Ogston, W. D. 1966. Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease* 143:338-347.
- Condon, W. S., and Ogston, W. D. 1967. A segmentation of behavior. *Journal of Psychiatric Research* 5:221-235.
- Dittman, A. T., and Llewellyn, Lynn A. 1969. Body movement and speech rhythm in social conversation. *Journal of Personality and Social Psychology* 11:98-106.
- Duncan, Starkey, Jr. 1969. Nonverbal communication. *Psychological Bulletin* 72:118-137.
- Efron, D. 1942. *Gesture and environment*. New York: Kings Crown Press.
- Eibl-Eibesfeldt, I. 1970. *Ethology: The biology of behavior*. New York: Holt, Rinehart and Winston.
- Ekman, Paul. 1971. Universals and cultural differences in facial expressions of emotions. To appear in *Nebraska Symposium on Motivation*, 1972, ed. J. Cole. Lincoln: University of Nebraska Press. [Read in preprint.]
- Ekman, Paul, and Friesen, W. V. 1969. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica* 1:49-98.
- Ekman, Paul, Friesen, W. V., and Tomkins, S. S. 1971. Facial Affect Scoring technique (FAST): A first validity study. *Semiotica* 3:37-58.
- Ex, J., and Kendon, A. 1964. A notation for facial postures and bodily

- positions in the social performance. Appendix to appendix II of progress report to Department of Scientific and Industrial Research for the Social Skills Project, Institute of Experimental Psychology, Oxford University, October, 1964. [Reprinted in *Social interaction*, by M. Argyle. New York: Atherton Press, 1970.]
- Firth, J. R. 1957. *Papers in linguistics, 1934-1951*. London: Oxford University Press.
- Goffman, E. 1963. *Behavior in public places*. New York: The Free Press of Glencoe.
- Goffman, E. 1971. *Relations in public*. New York: Basic Books.
- Grant, E. C. 1968. An ethological description of non-verbal behaviour during interviews. *British Journal of Medical Psychology* 41:177-184.
- Grant, E. C. 1969. Human facial expression. *Man* NS4:525-536.
- Hall, E. T. 1959. *The silent language*. Garden City, N. Y.: Doubleday.
- Hall, E. T. 1966. *The hidden dimension*. Garden City, N. Y.: Doubleday.
- Hoof, J. A. van. 1972. A comparative approach to the phylogeny of laughter and smiling. In *Aspecten van her Sociale Gedrag en de Communicatie bij Humane en Hogre Niet-Humane Primaten*, part 2. Rotterdam. [Also to appear in *Nonverbal communication*, ed. R. Hinde. London: Cambridge University Press.]
- Hoof, J. A. van. 1962. Facial expressions in higher primates. *Symposia of the Zoological Society of London* 8:97-125.
- Kendon, A. 1972. Some relationships between body motion and speech. In *Studies in dyadic communication*, ed. A. Seigman and B. Pope. Elmsford, N. Y.: Pergamon Press.
- Malinowski, B. 1923. The problem of meaning in primitive languages. In *The meaning of meaning*, by C. K. Ogden and I. A. Richards, appendix 3. London: Routledge and Kegan Paul.
- Mandelbaum, D. G., ed. 1963. *Selected writings of Edward Sapir in language, culture, and personality*. Berkeley and Los Angeles: University of California Press.
- Pike, K. L. 1967. *Language in relation to a unified theory of the structure of human behavior*. 2d ed. The Hague: Mouton.
- Reece, M. M., and Whitman, N. R. 1962. Expressive movement, warmth, and verbal reinforcement. *Journal of Abnormal and Social Psychology* 64:234-236.
- Reusch, J., and Bateson, Gregory. 1951. *Communication: The social matrix of psychiatry*. New York: W. W. Norton.
- Rosenfeld, H. M. 1966. Approval-seeking and approval-inducing functions of verbal and non-verbal responses in one dyad. *Journal of Personality and Social Psychology* 4:597-605.
- Schefflen, A. E. 1964. The significance of posture in communication systems. *Psychiatry* 27:316-331.
- Schefflen, A. E. 1965. *Stream and structure of communicational behaviour*. Philadelphia: Eastern Pennsylvania Psychiatric Institute. [To be republished by Indiana University Press, Bloomington.]
- Vine, Ian. 1970. Communication by facial-visual signals. In *Social behaviour in birds and mammals: Essays on the social ethology of animals and men*, ed. J. Crook. London: Academic Press.
- Waldron, R. A. 1967. *Sense and sense development*. London: Andre Deutsch.

Entropy and Art: An Essay on Disorder and Order

By Rudolph Arnheim. Berkeley and Los Angeles: University of California Press, 1971. Pp. 64. \$4.50.

This is a very short essay in which the author discusses such concepts as entropy, information, order and probability, homeostasis, tension reduction, art, and structure. The main thesis is that humans strive toward order and that the existing interpretations of tension reduction, chance, and the second law of thermodynamics do not provide a basis for understanding this motive. By now, this is a familiar argument, one which has been made by various theorists in the last several decades (Schrodinger, for example). The particular case against tension reduction and for some type of need for stimulation is even more familiar, yet no reference to the extensive literature on this point (Berlyne, Butler, Dember, Fiske and Maddi, Harlow, Hunt, Walker, or others) is made.

A central concern of the essay is with structure, a concept which remains elusive throughout, since no real attempt is made to identify what the term refers to, although an occasional example of a gestaltlike principle of organization is given. As for the structure of visual perception, again, almost no reference is made to research since 1950. Nor is there any clear presentation of exactly what is meant by entropy, information, and so on, which occasionally leads to serious ambiguity as to Arnheim's meaning.

If the essay's purpose is to discuss the issues in a somewhat dated but philosophically interesting and readable manner, so as to stimulate interest in the author's main thesis, then it succeeds. But if its purpose is to add to our understanding of these issues, within the framework of the very relevant scientific contributions of the last 20 years, then it fails.

Paul C. Vitz, *New York University*

The Measurement of Sensation: A Critique of Perceptual Psychophysics

By C. Wade Savage. Berkeley and Los Angeles: University of California Press, 1970. Pp. 578. \$15.00.

When one picks up a book entitled *The Measurement of Sensation: A Critique of Perceptual Psychophysics*, one might be expecting a volume that concerns itself with methodological issues drawn from the contemporary measurement morass which psychophysicists and perceptual psychologists find themselves continually facing during the course of empirical investigations of perceptual and sensory processes. This view would be further reinforced by a glance at the chapter and section headings. Once one begins reading, however, it becomes immediately apparent that this is not the focus of the book.

Wade Savage is a philosopher, and although he seems to be addressing himself to psychological issues in perception, he is really offering a philo-

sophical treatise on the theory of measurement applied to the measurement of sensory magnitudes. The first portion of the book uses loudness and pitch as examples and ends up concluding that no measurement of sensory magnitude is possible at all (something which all of us who work in perceptual laboratories are willing to admit about two or three times a week). He bases this conclusion on several different definitions of an adequate measuring scale, including those of Campbell and Stevens. However, he offers a set of alternatives based on the comparison of perceived magnitude of the test stimulus against the perceived magnitude of a stimulus designated as having unit intensity. His measure of loudness seems to reduce to a variant of the ascending method of limits, and his measure of pitch seems to be a variant of scaling by summing just noticeable differences, only here instead of *jnds* one uses halftone increments. Neither of these procedures strikes the reader as a major methodological breakthrough nor as a significant refinement of extant procedures.

Chapters 8 and 9 are probably the most interesting, at least to psychologists. Chapter 8 presents Fechner's method of scaling based on the discrimination threshold. The critical evaluation of this method includes all of the traditional arguments against the validity of the logarithmic sensory-scaling function, beginning with Fechner's assumption that *jnds* are additive. In chapter 9 Stevens receives similar treatment. This chapter is basically an updated version of Savage's 1966 paper in *Psychological Monographs*. Magnitude estimation is treated as nothing more than introspection with numeric instead of nominal categories. These chapters are occasionally interesting, but the rather involved reasoning, based frequently on semantics rather than content, will probably do little to shake the faith of the adherents of either form of scaling. Unfortunately the discussion of actual scaling procedures ends with chapter 9, providing no consideration of any of the more modern psychophysical measurement procedures.

The final chapters include a complex discussion on the conceptual difficulties associated with the distinction between physical and psychological dimensions. In sum, then, this is a book that despite the promise of the title really belongs on the shelf of a philosopher, next to his copy of Campbell's treatise on measurement theory, rather than on the shelf of a perceptual psychologist, sandwiched between his copy of Helmholtz and a battered edition of Stevens's handbook.

Stanley Coren, *New School for Social Research*

Fundamentals of Experimental Psychology

By Charles L. Sheridan. New York: Holt, Rinehart and Winston, 1971. Pp. xii, 402.

Designed for a first course in experimental psychology, this volume has six chapters—two on method and four surveying selected areas of psychological research. The chapters on method range over a variety of topics, including broad aspects of philosophy of science, behaviorism, problems of inductive inference, the nature of statistical reasoning, and several aspects

of experimental design. The presentation is eclectic, drawing both on classical formulations and on more recent treatments emphasizing experimental control. The four research areas represented and the major topics covered are learning (operant and respondent conditioning, continuity versus non-continuity theories, insight versus trial and error, all-or-none versus incremental interpretations of verbal learning), and perception (classical psychophysical methods, sensory scaling, theory of signal detection, nativism versus empiricism, pain, color vision), and physiological psychology (Lashley's work on cortical functioning, sleep and dreaming, emotion), and complex human behavior (obesity, obedience, bystander apathy).

While the author's selection of topics seems reasonable, it remains essentially arbitrary. Of his selection, he writes that "the material included is tactically valuable, interests me, and has worked well with my students" (p. ix). In places, the organization could have been improved. For example, a section entitled "Some Commonly Used Experimental Designs" contains coordinate subsections on randomized designs, purposes of randomization, matched-subjects designs, repeated-measures designs, techniques for minimizing order effects, factorial designs, and yoked controls; the two subheadings under "Deliberate Confounding" are "Explanation and Pseudo-explanation" and "S-R versus R-R Relationships." Overall, however, the book's strengths outweigh its weaknesses. The author treats his subject thoughtfully, with authority, and in a balanced way. His presentation is literate and readable. Perhaps equally important, the reader senses the author's regard for his subject matter and is made to feel the material merits his serious consideration.

Kenneth H. Kurtz, *State University of New York at Buffalo*

Tactics of Social Influence

By Albert Mehrabian. Series in Clinical and Social Psychology and Personality. Englewood Cliffs, N. J.: Prentice-Hall, 1970. Pp. 143.

This volume describes in succinct form several major techniques for altering the behavior of others and oneself. The intended audience for the volume includes both the undergraduate student of psychology and the intelligent layman. Primary attention is given to the application of central concepts of behavior modification to a wide range of real-life concerns. The utilization of reward is extensively elaborated, while stimulus control, reciprocal inhibition, and punishment receive secondary attention. Also included are several techniques of influence not typically found among the wares of behavior modification. Modeling is considered an important adjunct to behavior-modification techniques, as are verbal requests for change and rehearsal. Strangely, the treatment includes almost no reference to the vast literature on social influence that has been developed in social psychology.

The volume is highly effective in challenging the reader to forgo his usual habits of 'muddling through' the problem of changing others and himself. It continuously stresses the advantages of systematic appraisal and analysis of complex situations, and provides much wisdom on alternative methods of change. However, the humanist will find repugnant the continued treatment

of man as an object of manipulation, and could well counter with his argument that concrete situations are so immensely complex and in such a state of continuous flux that systematic analysis can only provide a distorted and misleading simulacrum of social reality. From this viewpoint, abstract analyses of ongoing relations may best be used as sensitizing devices. Myopia would result from the fully dedicated application of behavior-modification principles.

Kenneth J. Gergen, *Swarthmore College*

Books received

(The books listed here have not as yet been noted in our pages. Listing here does not preclude their later review, however.)

- Aaker, David A., ed. *Multivariate Analysis in Marketing: Theory and Application*. Belmont, Calif.: Wadsworth Publishing, 1971. Pp. ix, 358.
- Agel, Jerome. *The Radical Therapist*. New York: Ballantine Books, 1971. Pp. 292. Paperback, \$1.25.
- Aiken, Michael, and Mott, Paul E., eds. *The Structure of Community Power*. Westminster, Md.: Random House, 1971. Pp. xi, 540.
- Anderson, David C. *Children of Special Value: Interracial Adoption in America*. New York: St. Martin's Press, 1971. Pp. 184. \$6.95.
- Andreas, Burton G. *Experimental Psychology*. 2d ed. New York: Wiley, 1972. Pp. xiv, 608.
- Astin, Alexander W. *Predicting Academic Performance in College*. New York: Macmillan, 1971. Pp. 299.
- Banaka, William H. *Training in Depth Interviewing*. New York: Harper & Row, 1971. Pp. x, 196. Paperback, \$3.95.
- Barbara, Dominick A. *How to Make People Listen to You*. Springfield, Ill.: Charles C Thomas, 1971. Pp. 180. \$6.75.
- Barber, Theodore; DiCara, Leo V.; Kamiya, Joe; Miller, Neal E.; Shapiro, David; and Stoyva, Johann. *Biofeedback and Self-control*. Chicago: Aldine-Atherton, 1972. Pp. xxi, 806. \$17.50.
- Barrett, James H. *Gerontological Psychology*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xv, 165. \$8.00.
- Becker, Ernest. *The Birth and Death of Meaning: An Interdisciplinary Perspective on the Problem of Man*. 2d ed. New York: The Free Press, 1971. Pp. x, 228. \$6.95.
- Bloch, Rene. *Bild und Personlichkeit*. Bern: Hans Huber, 1971.
- Boden, Margaret A. *Purposive Explanation in Psychology*. Cambridge, Mass.: Harvard University Press, 1972. Pp. 408, \$13.50.
- Branden, Nathaniel. *The Disowned Self*. Los Angeles: Nash Publishing, 1972. Pp. 256. \$7.95.
- Brown, F. Christine. *Hallucinogenic Drugs*. Springfield, Ill.: Charles C Thomas, 1972. Pp. ix, 154. \$10.50.

- Bry, Adelaide, ed. *Inside Psychotherapy*. New York: Basic Books, 1972. Pp. xxviii, 162. \$6.95.
- Burr, Helen Turner, ed. *Psychological Functioning of Older People*. 3d ed. Springfield, Ill.: Charles C Thomas, 1971. Pp. xviii, 108. \$7.50; paperback, \$4.50.
- Burton, Arthur. *Interpersonal Psychotherapy*. Englewood Cliffs, N. J.: Prentice-Hall, 1972. Pp. xv, 172. \$8.95.
- Bzoch, Kenneth R., and League, Richard. *Assessing Language Skills in Infancy: A Handbook for the Multidimensional Analysis of Emergent Language*. Gainesville, Fla.: Tree of Life Press, 1971. Pp. 56.
- Cain, William, and Marks, Lawrence. *Stimulus and Sensation: Readings in Sensory Psychology*. Boston: Little, Brown and Co., 1971. Pp. 325. Paperback, \$5.95.
- Cancro, Robert, ed. *Intelligence: Genetic and Environmental Influences*. New York: Grune & Stratton, 1971. Pp. xii, 312. \$12.50.
- Carkhuff, Robert R. *The Development of Human Resources: Education, Psychology, and Social Change*. New York: Holt, Rinehart and Winston, 1971. Pp. xviii, 422.
- Cattell, Raymond B. *Abilities: Their Structure, Growth, and Action*. Boston: Houghton Mifflin, 1971. Pp. xxii, 583. \$17.50.
- Cattier, Michel. *The Life and Work of Wilhelm Reich*. New York: Horizon Press, 1972. Pp. 224. \$6.95.
- Chess, Stella; Korn, Sam J.; and Fernandez, Paulina B. *Psychiatric Disorders of Children with Congenital Rubella*. New York: Brunner/Mazel, 1971. Pp. xi, 178. \$7.50.
- Chomsky, Carol. *The Acquisition of Syntax in Children from Five to Ten*. Cambridge, Mass.: The M.I.T. Press, 1971. Pp. 126. Paperback, \$1.95.
- Colbert, John, and Hohn, Marcia. *Guide to Manpower Training*. New York: Behavioral Publications, 1971. Pp. xi, 112. \$8.95.
- Cole, Michael; Gay, John; Glick, Joseph, A.; and Sharp, Donald W. *The Cultural Context of Learning and Thinking: An Exploration in Experimental Anthropology*. New York: Basic Books, 1971. Pp. xx, 304. \$10.00.
- Colquhoun, W. P., ed. *Biological Rhythms and Human Performance*. New York: Academic Press, 1971. Pp. x, 279. \$14.50.
- Connell, R. W. *The Child's Construction of Politics*. Portland, Oreg.: Melbourne University Press, 1971. Pp. 251. Paperback, \$4.55.
- Cortes, Juan B., and Gatti, Florence M. *Delinquency and Crime: A Biopsychosocial Approach*. New York: Seminar Press, 1972. Pp. x, 468. \$14.50.
- Curran, Charles A. *Psychological Dynamics in Religious Living*. New York: Herder and Herder, 1971. Pp. 288. \$6.95.
- Davidson, Alex. *The Returns of Love: A Contemporary Christian View of Homosexuality*. Downers Grove, Ill.: Intervarsity Press, 1971. Pp. 93. Paperback.
- DeVito, Joseph A. *Psycholinguistics*. Indianapolis: Bobbs-Merrill, 1971. Pp. 36. Paperback, \$.95.
- DeVries, Walter, and Tarrance, V. Lance. *The Ticket-Splitter: A New Force in American Politics*. Grand Rapids: Wm. B. Eerdmans, 1972. Pp. 149. \$4.95.

- DeWaele, J. P. *La methode des cas programmes en criminologie*. Brussels: Charles Dessart, 1971. Pp. 218. Fr45.40.
- Diethelm, Oskar. *Medical Dissertations of Psychiatric Interest (Printed before 1750)*. Basel: S. Karger, 1971. Pp. viii, 211. \$10.00.
- Dodwell, Peter C. *Perceptual Processing: Stimulus-Equivalence and Pattern Recognition*. New York: Appleton-Century-Crofts, 1971. Pp. xiii, 535. \$14.95.
- Eibl-Eibesfeldt, Irenaus. *Love and Hate: The Natural History of Behavior Patterns*. New York: Holt, Rinehart and Winston, 1972. Pp. 276. \$7.95.
- Eissler, K. R. *Talent and Genius: The Fictitious Case of Tausk contra Freud*. Chicago: Quadrangle Books, 1971. Pp. xii, 403. \$12.95.
- Evans, Wayne, O., and Kline, Nathan S., eds. *Psychotropic Drugs in the Year 2000: Use by Normal Humans*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xxiii, 168. \$7.75.
- Eyres, Alfred, and Dorn, Alfred. *Stress and Distress: A Psychiatrist's Guide to Modern Living*. Cranbury, N. J.: A. S. Barnes, 1971. Pp. 291. \$6.95.
- Fassler, Joan. *Don't Worry Dear*. New York: Behavioral Publications, 1971. \$3.95.
- Fassler, Joan. *My Grandpa Died Today*. New York: Behavioral Publications, 1971. \$3.95.
- Fassler, Joan. *The Boy with a Problem*. New York: Behavioral Publications, 1971. \$3.95.
- Fernandez, Ronald. *Social Psychology through Literature*. New York: Wiley, 1972. Pp. xviii, 477. Paperback.
- Fodor, Nandor. *Freud, Jung, and Occultism*. New York: Lyle Stuart, 1972. Pp. 272. \$7.95.
- Ford, Peter S. *The Healing Trinity: Prescriptions for Body, Mind, and Spirit*. New York: Harper & Row, 1971. Pp. xv, 130. \$5.95.
- Fraisse, Paul, and Piaget, Jean, eds. *Experimental Psychology: Its Scope and Method*. Vol. 4, *Learning and Memory*. New York: Basic Books, 1971. Pp. viii, 376. \$12.50.
- Friedland, William H., and Nelkin, Dorothy. *Migrant: Agricultural Workers in America's Northeast*. New York: Holt, Rinehart and Winston, 1971. Pp. xix, 281. Paperback.
- Gay, Dennis A.; Reagles, Kenneth W.; and Wright, George N. *Rehabilitation Client Sustention: A Longitudinal Study*. Madison: University of Wisconsin Regional Rehabilitation Research Institute, 1971. Pp. xiv, 60.
- Goldfried, Marvin R.; Sticker, George; and Weiner, Irving G. *Rorschach Handbook of Clinical and Research Applications*. Englewood Cliffs, N. J.: Prentice-Hall, 1971. Pp. xii, 436. \$14.95.
- Goodacre, Elizabeth J. *Children and Learning to Read*. London: Routledge and Kegan Paul, 1972. Pp. ix 117. \$3.25; paperback, \$1.75.
- Gordon, George N. *Persuasion: The Theory and Practice of Manipulative Communication*. New York: Hastings House, 1971. Pp. xvii, 558. \$16.50; paperback, \$10.
- Greenbaum, Richard, and Tilker, Harvey A., eds. *The Challenge of Psychology*. Englewood Cliffs, N. J.: Prentice-Hall, 1972. Pp. xvi, 505. \$6.95.

- Gregory, R. L. *The Intelligent Eye*. New York: McGraw-Hill, 1971. Pp. 191. Paperback, \$2.95.
- Hale, Nathan G., Jr. *Freud and the Americans: The Beginnings of Psychoanalysis in the United States, 1876-1917*. New York: Oxford University Press, 1971. Pp. xvi, 574. \$15.00.
- Halleck, Seymour, L. *Psychiatry and the Dilemmas of Crime: A Study of Causes, Punishment, and Treatment*. Berkeley and Los Angeles: University of California Press, 1971. Pp. xiv, 382. Paperback, \$3.95.
- Hashimoto, K.; Kogi, K.; and Grandjean E. *Methodology in Human Fatigue Assessment*. London: Taylor & Francis, 1971. Pp. xxx, 200.
- Henle, Mary, ed. *The Selected Papers of Wolfgang Kohler*. New York: Live-right, 1971. Pp. x, 465. Paperback, \$3.95.
- Henry, Jules. *Pathways to Madness*. New York: Random House, 1972. Pp. xxii, 477. \$10.00.
- Horrocks, John E., and Jackson, Dorothy W. *Self and Role: A Theory of Self-process and Role Behavior*. Boston: Houghton Mifflin, 1972. Pp. viii, 227. \$5.95.
- Howells, John G., ed. *Modern Perspectives in Adolescent Psychiatry*. New York: Brunner/Mazel, 1971. Pp. xxii, 614. \$20.00.
- Hunt, J. McVicker, ed. *Human Intelligence*. New Brunswick, N. J.: Transaction Books, 1972. Pp. 283. \$6.50; paperback, \$2.95.
- Hyder, O. Quentin. *The Christian's Handbook of Psychiatry*. Old Tappan, N. J.: Fleming H. Revell, 1971. Pp. 192. \$4.95.
- Jeffery, C. Ray. *Crime Prevention through Environmental Design*. Beverly Hills, Calif.: Sage Publications, 1971. Pp. 290. \$10.00.
- Johnson, Donald M. *Systematic Introduction to the Psychology of Thinking*. New York: McGraw-Hill, 1972. Pp. x, 498.
- Johnson, Homer, H., and Solso, Robert L. *An Introduction to Experimental Design in Psychology: A Case Approach*. Scranton: Harper & Row, 1971. Pp. viii, 216. Paperback, \$3.95.
- Julesz, Bela. *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press, 1971. Pp. xiv, 406. \$20.
- Jung, John. *The Experimenter's Dilemma*. New York: Harper & Row, 1971. Pp. x, 306. \$4.95.
- Kahn, Jack H. *Human Growth and the Development of Personality*. 2d ed. Oxford: Pergamon Press, 1971. Pp. xx, 242.
- Kaplan, Berton H.; Leighton, Alexander; Murphy, Jane M.; and Freyberg, Nicholas. *Psychiatric Disorder and the Urban Environment*. New York: Behavioral Publications, 1971. Pp. x, 310. \$12.05.
- Katz, Sanford N. *When Parents Fail: The Law's Response to Family Break-down*. Boston: Beacon Press, 1971. Pp. xv, 251. \$12.50.
- Katz, Solly; Wright, George N.; Reagles, Kenneth W. *The Impact of an Expanded Vocational Rehabilitation Program upon Intra-Agency Processes and Procedures*. Madison: University of Wisconsin Regional Rehabilitation Research Institute, 1971. Pp. xiv, 56.
- Kendler, Howard H., and Spence, Janet T., eds. *Essays in Neobehaviorism: A Memorial Volume to Kenneth W. Spence*. New York: Appleton-Century-Crofts, 1971. Pp. 345. \$12.45.

ANNOUNCEMENT

Proceedings of the Sixty-fourth Annual Meeting of the Southern Society for Philosophy and Psychology

The sixty-fourth annual meeting of the Southern Society for Philosophy and Psychology was held on March 30—April 1, 1972, at the Sheraton-Jefferson Hotel in Saint Louis, Missouri. There were 179 members, friends, and guests in attendance. The program consisted of 41 papers in philosophy and 65 in psychology, including three symposia in psychology, and three joint sessions. There was also a festschrift program in honor of Professor Harold Lee. As usual, a social hour and business meeting were included in the program.

At the first of the three joint sessions, Glenn Hawkes of Virginia Commonwealth University chaired a symposium on "Skinner's *Beyond Freedom and Dignity*"; E. Maynard Adams (University of North Carolina), Robert G. Burton (University of Georgia), Earl A. Alluisi (University of Louisville), and William S. Verplank (University of Tennessee) participated. At the second joint session, past-president Van de Vate introduced Stanley Williams, College of William and Mary, who delivered the presidential address, "Psychology's First 100 Years." John Compton of Vanderbilt University chaired the third joint session, in which Stuart Cook of the University of Colorado delivered an invited address, "Ethical Principles for Research with Human Subjects"; discussants were David H. Jones (College of William and Mary) and John Marshall (University of Virginia).

Two Richard M. Griffith Memorial Awards were given. The recipient in philosophy was Professor Allen Gibbons of East Carolina University, for his paper "Pseudo-Particulars." The recipient in psychology was Professor L. James Shapiro of the University of Manitoba, for his paper "Imprinting: Another Look at the Cheshire Cat." These awards are given annually for the papers of highest merit presented by members who have not yet received the doctoral degree or have received it within the last five years.

Cochairmen of the committee on local arrangements were Donald L. Sievert (philosophy) and Marion E. Bunch (psychology), both of Washington University. Cochairmen of the program committee were Thomas H. Hearn (philosophy, College of William and Mary) and Glenn R. Hawkes (psychology, Virginia Commonwealth University). Cochairmen of the Richard M. Griffith Memorial Awards committee were Dwight Van de Vate (philosophy, University of Tennessee) and Malcolm Arnoult (psychology, Texas Christian University). Chairman of the exhibits committee was Fred Thumin of the University of Missouri. Cochairmen of the membership committee were Donald S. Lee (philosophy, Tulane University) and Lelon J. Peacock (psychology, University of Georgia). Roy Cebik of the University of Georgia and Charles Watson of Washington University served as co-chairmen of the placement service. At the annual business meeting, reports of the officers and of the several standing and ad hoc committees were received and accepted.

The deaths of members James A. Jordan, James E. Green, John P. Nafe,

Harold J. Ralston, Donald K. Adams, and Karl M. Dallenbach were noted with sorrow.

There were 21 philosophers and 27 psychologists elected to full or associate membership. Elected as *full members in philosophy* were Anne D. Adams, Joseph Binn, Peter Burkholder, James P. Franke, Joe H. Hicks, Martin M. Jones, Robert J. Mulvaney, Theodore D. Norhenhaug, Earl W. Redding, Norvin Richard, Donald Sievert, Donald H. Weiss, William E. Hoffman, Alexander von Schoenborn, David R. Hiley, and Bart F. Kennedy. Elected as *full members in psychology* were Richard Arvey, Gordon M. Burghardt, William H. Calhoun, Ralph L. Elkins, Robert M. Gagne, Susan M. Hereford, Alan Krasnoff, John M. Larsen, William M. Moon, Arnold Powell, Luis Proenza, Robert W. Schaeffer, Raymond R. Shrader, Jean H. Vanderplas, Donald T. Williams, Bert Mead, Karl P. Hunt, Jean Tracy, Vesta S. Gettys, Charles F. Gettys, and Franklin M. Berry.

Elected as *associate members in psychology* were Carl Eshbaugh, Henry J. Folse, Jr., Thomas M. Johnston, John B. McClatchey, Richard A. O'Neil, and Thomas Rockmore. Elected as *associate members in psychology* were Mary G. Boyd, Donald L. Corson, Odessa E. M. Hahn, Dale R. Meyer, Judson Mitcham, and Albert B. Ray.

The following officers were elected to serve during 1972-1973: Lelon J. Peacock, University of Georgia, president-elect; Douglas Browning, University of Texas, president; William T. Blackstone to the council, for philosophy; Earl Alluisi to the council, for psychology; and Michel Loeb, secretary of the society. Continuing members of the council include Herbert Kimmel, University of South Florida; Ramon Lemos, University of Miami; Donald Sherburne, Vanderbilt University; and Clyde Noble, University of Georgia.

The sixty-fifth annual meeting of the Southern Society for Philosophy and Psychology will be held on April 19-21, at the Regency-Hyatt House in Knoxville, Tennessee. It will be hosted by the departments of philosophy and psychology at the University of Tennessee.

W. T. Blackstone, Secretary; *University of Georgia*

ANNOUNCEMENT

The Sixty-Eighth Annual Meeting of the Society of Experimental Psychologists

The sixty-eighth annual meeting of the Society of Experimental Psychologists was held at Stanford University, Stanford, California, on April 7 and 8, 1972. Ernest R. Hilgard, chairman of the society for the year, presided at the business meeting and at the sessions in which scientific papers were presented.

The following members and fellows were present: Amsel, Anderson, Attneave, Bourne, Boynton, DeValois, Grant, Grice, Guilford, Helson, Hilgard, Jenkins, Kamin, Kendler, Kimble, Lawrence, Lindsley, Logan, Mackworth, Marquis, G. Miller, Morgan, Posner, Postman, Pribram, Riesen, Rosenzweig, Shepard, Solomon, Thomas, Tulving, Underwood.

William Bevan, Lyle Bourne, Frank Restle, Patrick Suppes, and Allan Wagner were elected to membership in the society.

Charter Member Karl M. Dallenbach and Fellow Clarence H. Graham died during the year. The present membership of the society is 79 members and 42 fellows.

The 1972 Warren Medal was awarded to George A. Miller, with the citation "His talent for experiment has moved psycholinguistics from arm-chair to laboratory."

The society accepted the invitation of the University of Minnesota to meet in Minneapolis, Minnesota, in 1973. James J. Jenkins was elected chairman for the year. The executive committee consists of the chairman, the secretary, Howard Kendler, and David Grant. The trustees are Frank Geldard, Daniel Lehrman, and the secretary.

Carl P. Duncan, Secretary; *Northwestern University*

ANNOUNCEMENT

New Citation Index for Social Sciences

A new index designed to improve access to information published in the social sciences and humanities journal literature has been announced by the Institute for Scientific Information (ISI).

Beginning in 1973, ISI will offer a *Social Sciences and Humanities Citation Index* covering articles, book reviews, and correspondence appearing in approximately 600 key journals in the fields of anthropology, community health, demography, economics, educational research, ethnic studies, geography, history, law, linguistics, management, marketing, political science, psychiatry, psychology, sociology, statistics, and urban planning/development.

Plans for the new index were disclosed by Melvin Weinstock, ISI Senior Information Scientist, at the recent conference on Access to Knowledge and Information in the Social Sciences and Humanities, sponsored by Queens College of the City University of New York. Weinstock said that the social sciences index will be patterned after ISI's widely used *Science Citation Index*[®]. He claimed that the new reference tool is the result of continuing requests from social scientists to apply citation indexing to the literature relevant to their work.

Citation indexing is based on the fact that current articles referencing the same earlier publication have a subject relationship not only with the earlier work but also with each other. To utilize this relationship, a citation index lists together all articles published during a specified time period that reference the same earlier publication. In effect, the earlier work becomes an indexing term by which someone who is familiar with its content may look up new articles on the same subject.

Weinstock said that "terminology in the social sciences and humanities is particularly imprecise and subject to rapid change. This makes it very difficult for traditional 'subject-heading' indexing systems to draw together related material. Citation indexing, however, lets the best qualified person — the author — index his own work through his reference citations."

Quoting from a study by Maurice B. Line, Director of England's National Central Library, Weinstock told the conference that the "majority of social scientists questioned said they would find a citation index very useful in their literature searching." Weinstock also said that ISI has been conducting its own research on the need for the new citation index and what characteristics it should have in terms of format, coverage, frequency of issuance, cost, etc. He said that the final configuration of the index will reflect the results of this research.

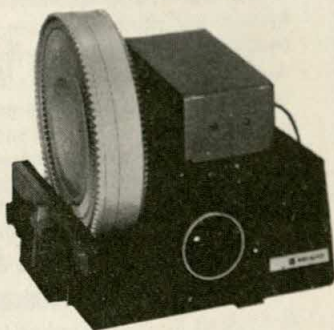
Anyone wishing more information or wanting to make suggestions on any aspect of the *Social Sciences and Humanities Citation Index* should call or write Melvin Weinstock, Institute for Scientific Information, 325 Chestnut Street, Philadelphia, Pennsylvania 19106.

NEW PHOTO-ELECTRIC PURSUIT ROTOR



The new Pursuit Rotor is manufactured in two models, the MK3 and the MK3/T which incorporate a digital 'time on target'. The variable speed range is from 5 to 50 revolutions per minute and the sampling periods are standardised on 5 — 10 and 20 seconds for simplicity of operation. The MK3/T, as depicted, has the above features and also includes a 4 digit totalising counter to eliminate the use of an external 'total time on target' counter.

SINGLE CHANNEL GROUP TACHISTOSCOPE



The instrument is a general purpose slide projector with remote control facilities and interchangeable lenses. The projector has been modified by the addition of a high-speed galvanometer which operates a shutter membrane to obscure the light path. The rise time of the shutter is 1.5 milli-seconds. This instrument can be used in word, picture and symbol recognition and subliminal research.



Forth Instruments Ltd.,
Engine Road, Loanhead, Midlothian,
Scotland.
Telephone: Loanhead 104.

Manufacturers of Single Channel Group Tachistoscopes, Pulse Generators, Memory Drums, Counter Timers, Voice Switches, Micro Electrode Pullers, Reaction Testers.

"In a moment you will fall deeply asleep. When you do, a dream will form. You will find yourself sitting in a movie theater looking at a blank screen. . . ."

With these words and careful suggestions from the therapist, the patient will fall into a hypnotic sleep and will redream the previous night's experience, in a form of psychotherapy examined in

DREAMS, IMAGES, AND FANTASY

A Semantic Differential Casebook

by C. Scott Moss, with a foreword by Charles E. Osgood

"Moss writes with considerable expertise on the subject of how hypnosis can effectively be used to facilitate psychotherapy that is intensive, time limited, and crisis oriented. He artfully combines the understanding gained from a highly regarded quantitative measurement of meaning, Osgood's semantic differential with his own clinical sensitivities, to illustrate the value of hypnosis in reaching and utilizing the symbol-translating mechanisms of the mind. This book is unique in the field. . . ." — *Choice*, \$8.50

"BLACK ROVER, COME OVER!"

The Hypnosymbolic Treatment of a Phobia

by C. Scott Moss and Pauline James

Published to accompany *Dreams, Images, and Fantasy*, this book stands on its own as an account of the swift and successful treatment of a patient with an extreme phobia. It "consists largely of James's personal account of her therapeutic experience, accompanied by brief notes by psychologist Moss of his interpretations of the same therapy sessions. . . . a fascinating, well written story." — *Library Journal*.
". . . a fascinating view of what may be accomplished in the alleviation of symptoms and the development of personal freedom when a brilliant, dedicated therapist cautiously utilizes hypnotic techniques to elucidate dream material." — *American Journal of Orthopsychiatry*. Paper, \$3.95. With audiotape of a transcribed therapy session (1 hour, 45 minutes, 3¾ IPS, open reel or cassette), \$12.50

UNIVERSITY OF ILLINOIS PRESS  Urbana Chicago London



Cornell
University
Press

Systematic Psychology

Prolegomena

By EDWARD BRADFORD TITCHENER. With a Foreword by Rand B. Evans and Robert B. MacLeod. Published posthumously in 1929, this classic work is still the most comprehensive analysis of non-behavioristic systems of psychology ever written. It is reissued with a Foreword and new notes by Rand B. Evans and Robert B. MacLeod, who place it in the context of modern systematic psychology and demonstrate its importance to understanding classical psychological systems.

\$7.50 cloth; \$2.45 paper.

Surface Color Perception

By JACOB BECK. This is the first comprehensive and systematic review of experimental and theoretical research on the problem of surface color perception. The author discusses the early studies of Katz, Kardos, and Katona, the studies at Columbia under Woodworth, those of Helson and Judd, and the more recent findings on sensory contrast and adaptation.

\$11.50

Abnormality and Normality

The Mothering of Thalidomide Children

By ETHEL ROSKIES. This case study of a highly unusual form of maternity is a valuable addition to the literature on handicapped or deviant children. It is an account of how mothers who took part in a government-sponsored habilitation program in Montreal perceived the process of bearing and rearing (or deciding not to rear) a child with congenital thalidomide-induced deformities. Professor Roskies traces how the biological, psychological, and social factors interacted—and changed over time.

\$12.50

Cornell University Press ITHACA AND LONDON

667
11-9-10
AKC
Docket 2 issue
to the Librarian
R
4/9/73

THE AMERICAN JOURNAL OF PSYCHOLOGY

Edited by

Lloyd G. Humphreys
William E. Kappauf

M. E. Bitterman
Benton J. Underwood

Karl M. Dallenbach: 1887-1971 (Rand B. Evans)	463
WALTER C. GOGEL. Scalar Perceptions with Binocular Cues of Distance	477
JOEL ZIMMERMAN, JOHN J. SHAUGHNESSY, and BENTON J. UNDERWOOD. The Role of Associations in Verbal-Discrimination Learning	499
JACK RICHARDSON and SARA K. STANTON. Some Effects of Learning to a Set of Components on Stimulus Selection	519
TERRY J. SPENCER. Backward Masking: Stimulus Energy and Positional Uncertainty	535
DONALD W. STILSON, KJELL HAASETH, CHARLES R. WALSMITH, ROBERT J. SCHNEIDER, and MICHAEL ROGERS. Marbe's Law, Stimulus Expectancy, and Response Competition	549

continued on back cover

Editors Lloyd G. Humphreys M. E. Bitterman
William E. Kappauf Benton J. Underwood

Book Review Editor Julian Hochberg

Consulting Editors William Epstein Robert Fox Wendell Garner
Lawrence E. Jones John Krauskopf Michael Rashotte Allan Wagner

Index Editor Rand B. Evans Necrology Editor E. R. Hilgard

The *American Journal of Psychology* was founded in the interest of general experimental psychology and publishes reports of original research, shorter notes and discussions, and pertinent book reviews. Manuscripts of articles and notes may be submitted to Drs. Humphreys or Kappauf (University of Illinois at Urbana-Champaign, Champaign, Illinois 61820), Bitterman (University of Hawaii, Honolulu, Hawaii 96822), or Underwood (Northwestern University, Evanston, Illinois 60201). Books intended for review should be sent to Dr. Carl P. Duncan (Northwestern University, Evanston, Illinois 60201), who will be the new Book Review Editor as of 1973.

Authors of articles may order offprints at the time they return galley proof. Other contributors are sent tear-sheet copies as available unless extras are ordered when manuscript is accepted.

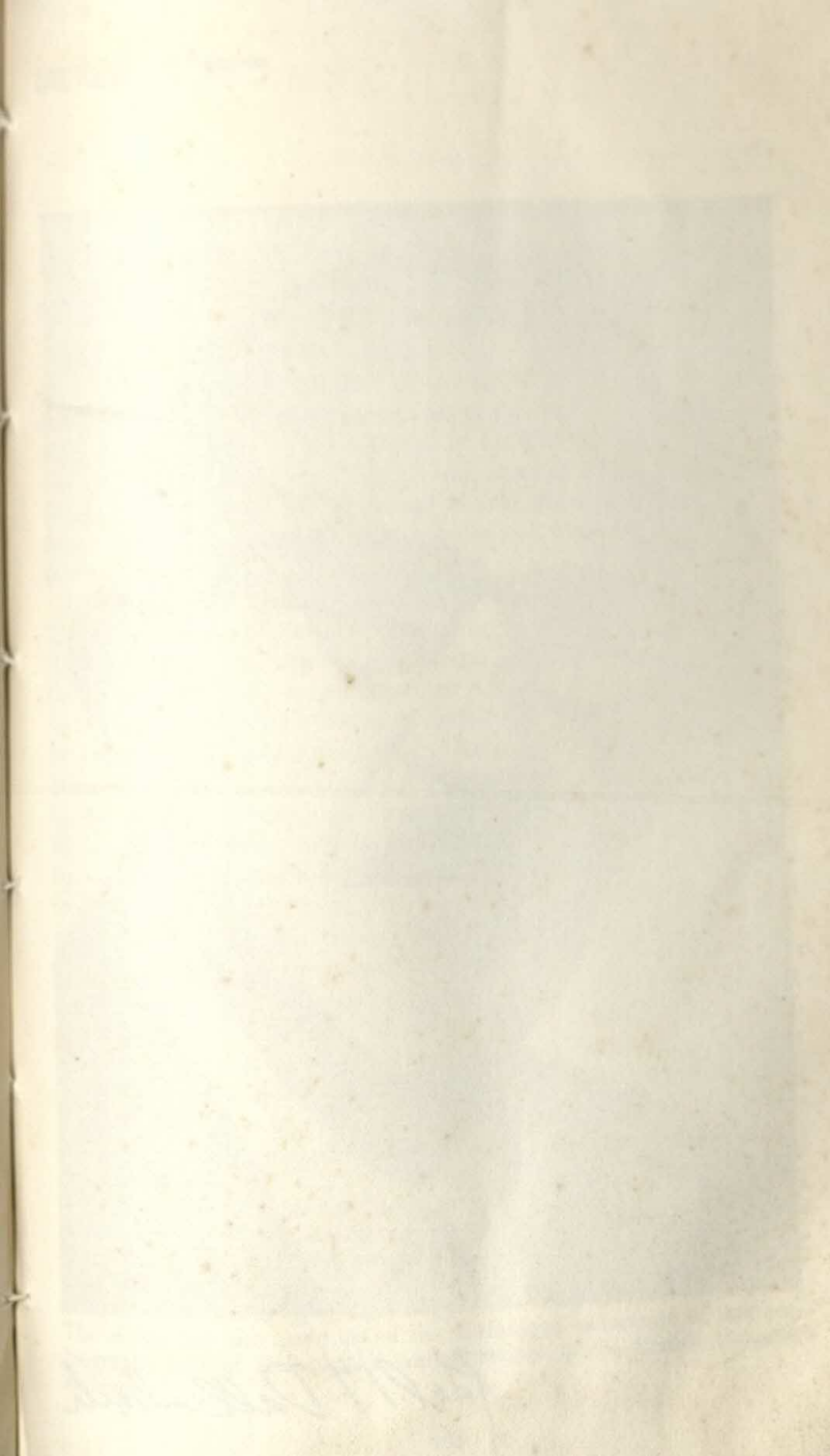
Contributors are asked to follow the style described in the *Publication Manual* of the American Psychological Association. Keep abstract and title quite brief, use footnotes only when essential, avoid all but the most common abbreviations (2 mm, 4 deg, 60 db, 15 sec; *but* subjects, experimenter, paired-associate learning), and use the tables and figures to supplement rather than repeat the text or each other; a well-designed table or figure can provide more information than many words of text. If you elect to omit from your manuscript any information that is critical to the evaluation of your data (analyses of variance summary tables, means, variances, correlation and factor matrices), provide two copies of that material for the editors' and their consultants' use in reviewing the paper.

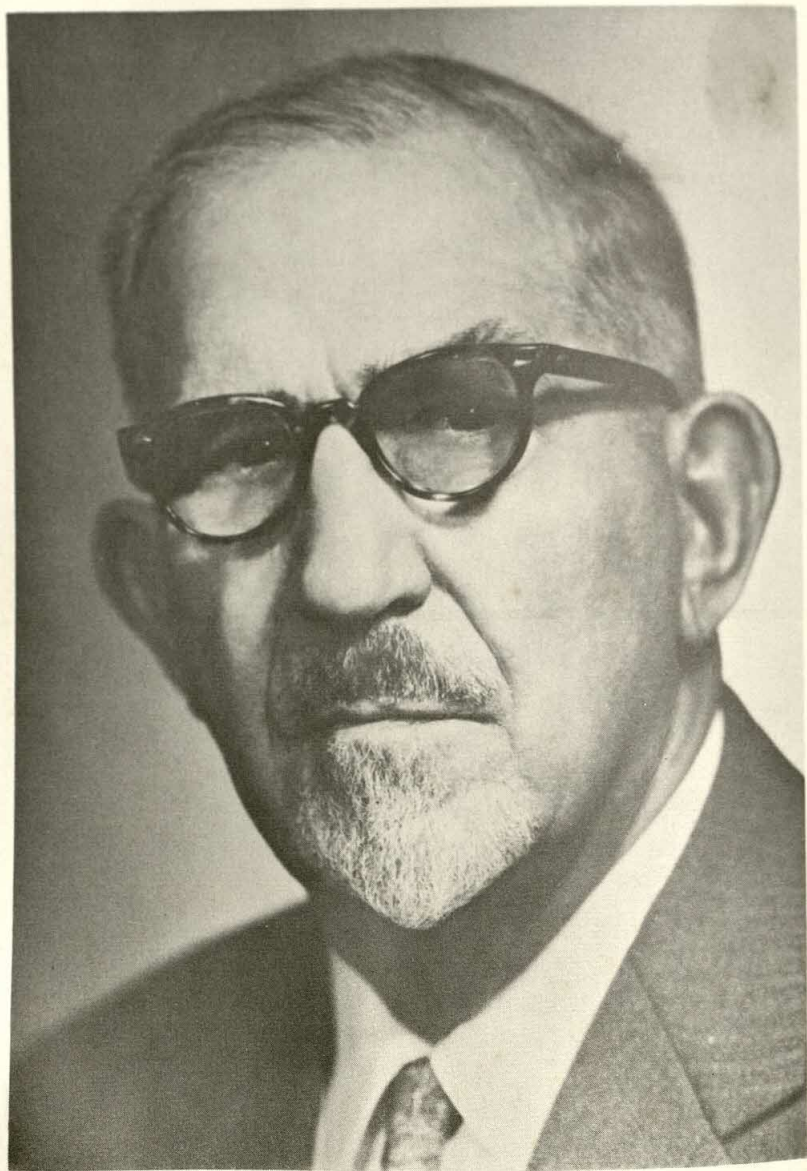
The subscription price is \$15.00 a year (foreign \$15.50); single copies cost \$5.00. Back numbers and volumes may be obtained from the University of Illinois Press, Urbana, Illinois 61801, to which all business inquiries (subscriptions, advertisements, changes of address) should be addressed.

Published quarterly, in March, June, September, and December
Office of publication, University of Illinois Press, Urbana, Illinois 61801
Second-class postage paid at Urbana, Illinois

11 SEP 1973

ROSEMARY COGAN and ALVIN G. GOLDSTEIN. Reporting of Fragmentations in the Binocular Rivalry of Contours	569
N. JACK KANAK, LAWRENCE E. COLE, and C. DWAYNE CURTIS. Retention-Task Carryover Effects after Verbal-Discrimination Learning	585
G. JAMES GOODWIN and DARRYL BRUCE. Temporal Dating and Single-Trial Free Recall	597
Book Reviews (J. E. R. Staddon, Stanley Coren, Robert R. Zimmerman, Stephen C. Woods; books received)	605
Contributors to Volume 85	619





Karl M. Dallenberg

KARL M. DALLENBACH: 1887-1971

The death of Karl M. Dallenbach on December 23, 1971, breaks one of the few remaining links this generation of psychologists has with the psychology of the first two decades of this century. His 60 years as a psychologist make the capsulization of his career impossible. He had the time and the interests for several careers, which brought him recognition as an authority in several areas of the science. One of those careers was his association of over 50 years with the AMERICAN JOURNAL OF PSYCHOLOGY, over 40 years as its editor.

Dallenbach was born on October 20, 1887, in Champaign, Illinois.¹ This was the year of Fechner's death and the year of the founding of the journal that was to play such a large part in his own life. Karl M. Dallenbach was the second of three sons born to John and Anna Mittendorf Dallenbach. His great-grandfather Dällenbach emigrated to America from Switzerland in the 1820s; the maternal grandfather Mittendorf came from Germany in the 1850s. He was always proud of his Swiss-German background and particularly that the name Dällenbach originated from the legendary Swiss patriot, Wilhelm Tell. Dällenbach means 'valley brook' in German and it was this name that he took for his *nom de plume*, Charles Valley Brook, in so many book reviews and subject indexes in the AMERICAN JOURNAL OF PSYCHOLOGY.

The matter of his name brought Dallenbach into confrontation with generations of university and civil bureaucrats. Although his mother seems to have intended the 'M.' in his name to stand for Mittendorf, it was never properly registered, so Dallenbach contended throughout his life that his middle name was 'M.,' including the period. Although Karl was his name, it was seldomly used except for formal signatures and was reserved primarily for his wife. To his parents he was, often

¹The author is indebted to Mrs. Ethel Douglas Dallenbach for her help and suggestions on this manuscript. More detailed coverage of various portions of Dr. Dallenbach's life may be found in E. G. Boring's article (Karl M. Dallenbach, this JOURNAL, 71, 1958, 1-40) and in Dallenbach's own autobiography (in *A History of Psychology in Autobiography*, 5, 1967, 59-91). The author has also made use of the unabridged manuscript of that autobiography and of Dallenbach's correspondence in the Cornell University archives.

as not, 'Charlemagne'; to his school friends, up until the time he went to Cornell, he was 'Dally'; to his close associates from the Cornell days on he was 'Dall'; and to his close students he was 'Dr. Dall.' He once received a letter from a psychologist he knew only vaguely addressing him as 'Karl.' "If he knew me well enough to call me 'Karl,' he would know enough *not* to call me 'Karl!'" he exclaimed.

Dallenbach's mother characterized him as having a "serious, sober, and lugubrious disposition" and attributed it to the month of his birth. He, however, blamed the description on his 'girlhood.' His mother expected her second child to be a girl. His curly hair tended to encourage her in this expectation, even though he was a boy. Until after his fourth birthday, he was dressed as a girl and was taught to play girls' games. Even after he was allowed masculine clothing, he was required to wear his hair in long curls. Only his refusal to go to school unless his hair was cut seems to have allowed him the appearance of a boy at last. His mode of dress and appearance during these early years gave him ample opportunity to learn self-defense. To the lessons learned in some of those early battles, he attributed his lifelong characteristics of aggressiveness and tenacity.

His first introduction to psychology was in 1907, in his second year at the University of Illinois. He was taking the pre-law curriculum, which required some psychology. He had avoided the course during his first year because he identified it with phrenology, but he found the course quite different from his expectations. The elementary course was conducted by John Wallace Baird, one of E. B. Titchener's prize students, and his lectures were patterned after those of the master. Dallenbach found Baird a fascinating lecturer and Titchener's *Outline of Psychology* "as interesting as a novel."² This pleasant introduction to psychology with Baird led him to elect other psychology courses, including Baird's in experimental psychology, where he came into contact with Titchener's *Manuals* for the first time.³

Baird was so impressed by this pre-law student's laboratory reports that he had them bound as examples for future students. In 1909, when Titchener later gave at Illinois his series of lectures on the thought processes, he also was impressed with the laboratory reports and suggested that Dallenbach come to Cornell for a doctorate in psychology. This was the first of the vectors that pushed Dallenbach into psychology.

He became Baird's assistant in the experimental psychology course

²E. B. Titchener, *An Outline of Psychology*, 3d ed., 1899.

³Titchener, *Experimental Psychology: A Manual of Laboratory Practice*, 1-4, 1901, 1905.

in 1908. After Baird left Illinois, however, Dallenbach's intentions toward psychology began to waver. It was only a fellowship in psychology at the University of Pittsburgh that kept him from going to Harvard and a degree in law.

It was during his freshman year at Illinois that he had first met Miss Ethel Douglas of Chicago. Sitting next to her in history class for the first time, he told her in a rather matter-of-fact way that he was going to marry her. "I think I shall have something to say about that," was her shocked reply. It took five years of persistence before he was to get anything but a cool response from her, and eight years before he convinced her that he was right, after all. This romance is a good example of Dallenbach's characteristic tenacity and determination.⁴

The year at Pittsburgh found law displaced by medicine as the primary competitor to psychology in his life. It was largely out of his loyalty to and admiration for Baird, and because Titchener had suggested Cornell, that he applied there for a fellowship in psychology. He received the Sage fellowship, and in the fall of 1911 found himself in Ithaca and in psychology. Even then, it was 1920 before the force of events snuffed out the ambition for the M.D.

This was the 'golden age' of Titchener's Cornell. Titchener's *Textbook* had appeared the year before, and the laboratory was buzzing with research activity.⁵ The Perky experiments and those of A. S. Edwards on the image element were being carried out.⁶ Dallenbach was one of the observers in Edwards' experiment that sealed the fate of the image as an elementary concept. E. G. Boring was there as Titchener's student assistant. It was the beginning of a long friendship between Dallenbach and Boring.

Titchener assigned Dallenbach the problem of attention and in particular the attempt to test Geissler's findings in modalities other than vision.⁷ In one of their discussions, Titchener discovered Dallenbach's ability to play blindfold chess. This interested Titchener mightily

⁴Dallenbach also displayed these qualities in sports, both at Illinois and at Pitt. He was particularly well known for his prowess in football, but he was also a prize-winner in checkers, chess, shot put, and hammer throw. Sports would always be a fascination to him, both on and off the field.

⁵Titchener, *A Textbook of Psychology*, 1910.

⁶C. W. Perky, An experimental study of imagination, this JOURNAL, 21, 1910, 422-452; A. S. Edwards, An experimental study of sensory suggestion, this JOURNAL, 26, 1915, 99-129.

⁷L. R. Geissler, The measurement of attention, this JOURNAL, 20, 1909, 473-529. Dallenbach's dissertation was published as The measurement of attention, this JOURNAL, 24, 1913, 465-507.

because it required imagery. These were still the days of the Titchener-Külpe battle over imageless thought. Titchener used the 1912 meeting of the Experimentalists to 'show off' Dallenbach's ability.⁸

It was Dallenbach who made the series of life masks of the psychology graduate students. The set hung in Titchener's office for many years. He also did a bas relief of Titchener after the style of Felix Pfeifer's casting of Wundt. Titchener posed for it and signed the original. Plaster casts were made for the members of the laboratory and a bronze casting was given to Titchener.

In the summer of 1912, Dallenbach went to Germany to hear Wundt's lectures. He planned the trip originally with a group of fellow students. One by one they dropped out, until finally he found himself on the dock alone. Undaunted, he set out on his own. On his arrival, however, he found that Wundt was away for the summer and was neither lecturing nor accepting visitors. By a stroke of fortune, the trip was saved when the U. S. consul suggested he go to Bonn and hear Oswald Külpe, the man Titchener had talked about so much. The contact he was able to have with Külpe gave Dallenbach a lifelong regard for the man. Before leaving for home, he did what none of the German students dared to do—to ask Külpe for his photograph. There was none available in the bookstores and he wanted one. His logic was typical: "The worst he could do was to say no, so why not ask?" Külpe was flattered at the request and Dallenbach got his photograph, much to the envy of the other students.

It was during this trip that Dallenbach's physiognomy gained the imperial beard. He did it as a surprise for his fellow students in the laboratory and fully intended to shave it off on his return. The response of his friends and particularly that of Miss Douglas was such that he kept it for the remainder of his life.

It was during that last busy year at Cornell that he first attended Titchener's lectures in elementary psychology. He would hear them many more times between then and 1925.

The Ph.D. was conferred in June, 1913, and Titchener sent him to the University of Oregon as an instructor. Sent is the correct word. It was still Dallenbach's intention to go back to Pitt and get his M.D. Titchener would not hear of it, so Dallenbach went to Oregon.

Oregon had a lovely campus and a highly social faculty. The problem was that there was no research activity. The head of the department was Edmund Conklin, a student of G. S. Hall. Conklin thought

⁸Dallenbach, Blindfold chess: The single game, in *Studies in Psychology: Titchener Commemorative Volume*, 1917, 214–230.

of his instructor as sort of a personal assistant—not Dallenbach's idea at all. This was an irritation, but it was the lack of impetus for research that bothered him the most. The experiment was what he had been prepared for and he did not quite know what to do without a laboratory. Not wishing to stagnate, even in so lovely an area, he left Oregon after his second year for the hopefully greener pastures of Ohio State.

It was during the Oregon period that he went to Chicago, where he proposed to Miss Douglas. She said yes and thus he kept his promise of eight years before. He often used her as the model for the kind of wife a psychologist should have—one who could understand the peculiarities of the academician and in particular one who did not have to be entertained continually and would help urge her husband on, even if it meant seeing him less. He was fortunate in his selection.

He had hardly settled at Ohio State when he got the call from Titchener to return to Cornell as an instructor. It meant a cut in pay, but Titchener expected a man to sacrifice to be at Cornell. To Cornell it was, and there Dallenbach would be for over 30 years, the first of his two academic careers.

He realized what going to Cornell meant in terms of his advancement. Psychology at Cornell had a definitely pyramidal structure. There were two full professorships, Titchener's and H. P. Weld's. He would come in at the bottom and would have to wait for someone to resign or die before he could move up to a higher notch. The road to the Sage professorship would be even longer and harder than he had imagined. But Cornell was where the research was and he immediately plunged into his work.

The entry of the United States into World War I would soon break into academic life at Cornell. Dallenbach wanted to do his part, and even though color blind, he managed to pass the eye test to get into the army. Actually, he only partly passed. He memorized the order of the color-test cards. This worked quite well for one eye, but something happened on the series on the second eye, with the result that the ophthalmologist believed he had stumbled onto a man with one normal and one color-defective eye. It took some talking to convince the ophthalmologist to wait until after he returned from the war before publishing anything on the 'phenomenon.' Dallenbach served in the newly organized Testing Corps of R. M. Yerkes, and for a time under Walter Hunter.

After his discharge from the army, he returned to Cornell. He was offered a job with a major trolley company at twice his Cornell salary. In those days, the trolley companies were employing a large number

of psychologists to do the early version of personnel testing. To him, of course, this was applied psychology and not science at all. He chose Cornell.

Things at Cornell were much as they had been when he left. The faculty still attended Titchener's elementary lectures and held the faculty meetings in Titchener's room behind the lecture hall. It was in one of those meetings that the matter of the AMERICAN JOURNAL OF PSYCHOLOGY came up. "The JOURNAL is going to hell!" Titchener exclaimed. G. Stanley Hall, now an old, irritable man, was letting things slide so far as the JOURNAL was concerned, and there was little that the associate editors, Titchener and E. C. Sanford, could do about it. "Why don't we buy it?" was Dallenbach's reply. To Titchener, who always had wanted a journal of his own (as had Wundt before him), the sharing of the AMERICAN JOURNAL had been unsatisfactory from the beginning. The idea of having the JOURNAL under his sole editorship was appealing. Dallenbach was given the job of looking into the matter.

The "we" in Dallenbach's suggestion was intended to be a financial partnership among the Cornell faculty, or failing that, a purchase by Cornell itself. It was in this belief that he bargained with Hall, billing himself as acting for an anonymous backer. He would soon learn how anonymous that backer was. Dallenbach bargained with Hall for several weeks and brought Hall's price down to half the original figure. This was still more than he had intended to pay, but the agreement was finally closed. Hall required a large deposit as 'earnest money' and Dallenbach paid it out of his own pocket. The deal being made, he returned to Ithaca to secure the remainder of the money.

Cornell, as it turned out, had no interest in acquiring the JOURNAL and the banks would not give a loan without security. The only security available in the Cornell department was Titchener's house and property. Titchener was unwilling to risk that, even for the JOURNAL. No one had the ready cash, so Dallenbach found himself his own 'anonymous backer.'

In order to save him the loss of his 'earnest money,' his father, a successful businessman, allowed him to borrow against his inheritance. His father believed the JOURNAL a poor financial risk, but the younger Dallenbach did not think of acquiring the JOURNAL as a financial move, of course. He did it partly because Titchener wanted it and partly because he found himself owning it without fully intending to.

Titchener was given full reign in all matters editorial and Dallenbach was business manager. Titchener demanded and received Dallenbach's

guarantee in writing that he would be allowed to be editor so long as he wished, but declined to give a similar guarantee that he would remain editor. Titchener seems not to have understood what Dallenbach's purchase of the *JOURNAL* meant in terms of ownership. Titchener was no businessman and had no appreciation for the legalities of property. When he did come to understand that the *JOURNAL* was privately owned—and by an 'underling'—as he did some five years later, it brought about his resignation and set the stage for something of a tragedy in both their lives.

It was during this period that Dallenbach began his 80-hour work-week, a habit he would maintain for a goodly portion of the rest of his life. The *JOURNAL* was a draining affair and required much more time than he had ever imagined, although his bibliography shows it did not decrease his experimental efforts.⁹

By the early 1920s he began to feel the psychological isolation that was Titchener's Cornell. There was an attempt on his part and Boring's to bring Titchener and Cornell, at least socially, into the mainstream of American psychology. They planned to do this by making Titchener president of the American Psychological Association. Titchener would have none of it, however, and he resigned his membership in the APA rather than be its president.

Dallenbach was determined not to be enveloped in the isolation of Ithaca and decided to break away from the anti-APA spirit of Cornell and work within the framework of the national society. It was in 1923 that he proposed to Boring and Samuel Fernberger that something like Titchener's Experimentalists be formed within the APA. This was intended to supplement the 'Spring Meetings' of Titchener's ever-narrowing group, not to replace it. As was his habit, Dallenbach went into the matter with a plan in hand but with a willingness to compromise—so long as his opponent was willing to work for his concessions. The result of his plan was the beginning of the Experimental Roundtable at the APA Christmas meetings. It was an informal gathering of experimental psychologists to talk on current research. It was much on Titchener's own plan but was designed for a wider variety of systematic creeds than Titchener's Experimentalists allowed. For many years this roundtable was to many the highlight of the APA meetings.

As early as 1923, definite strains began to show between Dallenbach and Titchener. Dallenbach, beginning to feel his maturity as a psychologist, did not always follow Titchener's rather rigid rules of etiquette.

⁹M. C. McGrade, A bibliography of the writings of Karl M. Dallenbach, this *JOURNAL*, 71, 1958, 41-44.

The first major 'blowup' occurred over as simple a matter as his planning a party for one of the men in the laboratory who was getting married—planning it without clearing the details with Titchener. There were other such disagreements, each one minor, but each adding to the strain in the relationship. Finally, in 1925, Titchener demanded that the JOURNAL be turned over to a holding company, thinking of Cornell as a likely organization. Dallenbach replied that he had no intention of giving up title to the JOURNAL. Titchener resigned immediately and broadcast his resignation to every major psychologist across the country. Dallenbach thus found himself with the oldest journal of psychology in the world and with its weight completely on his back. He turned to Boring, Madison Bentley, and Margaret Washburn, and with them formed a joint editorship. Titchener exclaimed that it was a group doomed to failure, since no one could keep those four together. In that statement, Titchener underestimated Dallenbach, who *did* hold the group together and successfully salvaged the JOURNAL from what Titchener assumed would be total destruction. One need only peruse the letters between Dallenbach and the coeditors, particularly those to the meticulous Boring, to see just how good and patient a manager he was. It was at this time that Charles Valley Brook was created, so that it would not look as though Dallenbach was doing too much on the JOURNAL.

The period between 1925 and 1927 was one of great stress. The pressures of the JOURNAL were much greater than before, and the pressures within Cornell were greater still. Titchener's bitterness toward him was unrelenting. By 1926, Dallenbach was determined to leave Cornell, and only Titchener's death from a brain tumor in 1927 prevented him. Dallenbach rationalized all of Titchener's behavior as being due to the effects of the tumor. The feeling of bitterness vanished and the old loyalties returned as before. His personal loyalty to Titchener and to his memory would never diminish.

The period of the 1920s was not all JOURNAL work, however. A faculty member of the period recalls that Dallenbach was the most productive researcher in the laboratory, and his bibliography bears this out. He continued his work on attention, begun with his thesis for Titchener. By the time he chose to lay the problem aside for other experimental topics, he was generally considered the authority on attention as introspectively considered.

It was also during this period that he accidentally cut a nerve in his index finger. The chance to replicate the experiments of Head and Boring was too good to pass up, so he and Harry Helson began what

must be the longest experiment in the history of psychology, tracing the return of sensitivity to the area.¹⁰ The experiment extended over 40 years, and there were plans with the present author for one last experimental sitting, a plan never realized because of Dallenbach's death.

It was also in the 1920s that his now-classic experiment with Jenkins on the effect of sleep versus activity on the retention of learned material was performed.¹¹ Also, throughout the late 1920s and early 1930s, he took his movie camera to the meetings of the Experimentalists, giving psychologists of today one of their rare glimpses of the psychologists of that period *in vivo*.

In 1930-31, Dallenbach was given a visiting professorship at Columbia. The offer of a permanent professorship there was soon made. Here was the chance for the full professorship that had not been forthcoming at Cornell. Believing that New York City was no place to bring up his three small children, he refused the offer.

After his return to Cornell from the Columbia stint, he found that Bentley, who was by then Titchener's successor as head of the department, had arranged for a new professorship to be created for him.

The 1930s were perhaps the most peaceful of his life. There was the JOURNAL, of course, and the difficulty of keeping his spirited coeditors pulling together, but there were none of the major strains of the 1920s. This decade was also one of research. There were many studies, but most dealt with the cutaneous senses and gustation. For his work on the temperature sense, he devised the temperature stimulator that still bears his name.¹² It was, until semiconductor technology came in, perhaps the best punctate source available for such experiments.

With the 1940s came World War II, and Dallenbach found his academic life interrupted for a second time. For half the decade he was part of the armed forces and also chairman of the Emergency Committee in Psychology of the National Research Council. He was justly proud of his work in both areas.¹³

The war claimed his son John, lost over the English Channel on his sixth mission, after D-Day. John was only a dissertation away from his doctorate in psychology. It was perhaps the greatest single tragedy

¹⁰Harry Helson, Autobiography, in *A History of Psychology in Autobiography*, 5, 1967, 202.

¹¹J. G. Jenkins and Karl M. Dallenbach, Obliviscence during sleep and waking, this JOURNAL, 35, 1924, 605-612.

¹²Dallenbach, Some new apparatus, this JOURNAL, 34, 1923, 90-94.

¹³Dallenbach, The emergency committee in psychology, National Research Council, this JOURNAL, 59, 1946, 496-582.

in Dr. Dallenbach's life, one that never diminished with time. He and Mrs. Dallenbach created scholarships at Illinois, Cornell, and Texas in John's memory.

Returning after the war, he found himself at last Sage Professor of Psychology at Cornell, Titchener's old chair. But neither at Cornell nor anywhere else was psychology the same as before the war. Applied and clinical psychology were the thing, and the pressures for it were great at Cornell. Those pressures and the political situation at Cornell made it obvious to him that the kind of department he wanted was not possible there. When he was offered the chairmanship of the Psychology Department at the University of Texas, as he was in 1948, he accepted. Here was a chance to build a department with an experimental core. It was the beginning of his second career in psychology. Here again, he went to Texas with a plan in hand—one that included a modern psychological laboratory building. He had to compromise on the points of applied and clinical psychology, but the experiment was still the core of the training of Texas students for the next decade.

The last few years at Cornell and the first decade at Texas found Dallenbach as involved as ever in research. The papers of this period are probably his best. Perhaps the most important of this work was that on 'facial vision.' The structure of the experiments is a model of scientific procedure. He was tackling the problem of how the blind are able to locate obstacles in their path without seeing them. The view then popular was that of facial vision—that the skin receptors of the face become sensitive to sound impressions. That view had the imprimatur of William James and was what the introspections of blind observers seemed to indicate. It was not a matter of skin receptors, of course, but audition, as Dallenbach's experiments with various co-authors demonstrated.¹⁴

It was also during the last few years at Cornell that he constructed his olfactorium, the first genuine *camera inodorata*.¹⁵ It was also during this period that he and Minami carried out their classic work on the effect of sleep and wakefulness on retention in the cockroach.¹⁶ It was

¹⁴M. Supa, M. Cotzin, and Karl M. Dallenbach, "Facial vision": The perception of obstacles by the blind, this JOURNAL, 57, 1944, 133-183; P. Worchel and Karl M. Dallenbach, "Facial vision": The perception of obstacles by the deaf-blind, this JOURNAL, 60, 1947, 502-553; Cotzin and Dallenbach, "Facial vision": The role of pitch and loudness in the perception of obstacles by the blind, this JOURNAL, 63, 1950, 485-515; C. H. Ammons, P. Worchel, and Karl M. Dallenbach, "Facial vision": The perception of obstacles out of doors by blindfolded and blindfolded-deafened subjects, this JOURNAL, 66, 1953, 519-553.

Dallenbach's only animal experiment, an extension of the Jenkins and Dallenbach study of the 1920s. He was always glad he did the experiment, if only to show that he *could* do an animal experiment if he wanted to.

The Texas decade (1948-1958) rivaled the 1920s as the busiest of his life. Besides the heavy administrative load, there was an overload of teaching that he took on voluntarily, to expand the curriculum. There was still the JOURNAL, of course, and the greater burden of editing, since Boring, the last of the 1925 board of cooperating editors, dropped by the wayside. For several years, Dallenbach was sole editor of the JOURNAL, then forming another editorial board in the mid 1950s with M. E. Bitterman (one of his students from Cornell days), E. B. Newman, and later, Leo Postman. Even with the new board, he still did the major share of the editorial work. It can be said without exaggeration that he and Mrs. Margaret McGrade, his secretary and business manager for the JOURNAL, were the JOURNAL in the 1950s and 1960s.

The year 1953 saw the opening of Mezes Hall, built to Dallenbach's specifications. Psychology at Texas had finally come into its own, its faculty having grown from 6 in 1948 to 22 in 1958, a massive expansion for those days.¹⁷

The editorial and administrative load finally began to take its toll in research, however. All of the experimental papers of this period were coauthored with students. His singly authored pieces were generally historical notes and criticism, such as his "Phrenology versus Psychoanalysis" and his "Place of Theory in Science," two of his best-received articles.¹⁸

In 1958, he stepped down as chairman of the department, since, according to university regulations, faculty had to retire or go on 'modified service' when they reached the age of 70. This sixth and final decade in academics was still a busy one. The editorial responsibilities of the JOURNAL weighed heavier and heavier upon him. There were still his courses on the history of psychology and on psychophysics and there were still students with whom to work.

His interests extended to learning again. The controversy concerning

¹⁵D. Foster, E. H. Scofield, and Karl M. Dallenbach, An olfactorium, this JOURNAL, 63, 1950, 431-440.

¹⁶H. Minami and Karl M. Dallenbach, The effect of activity upon learning and retention in the cockroach, this JOURNAL, 59, 1946, 1-58.

¹⁷Dallenbach, The psychological laboratory of the University of Texas, this JOURNAL, 66, 1953, 90-104.

¹⁸Dallenbach, Phrenology versus psychoanalysis, this JOURNAL, 68, 1955, 511-525; The place of theory in science, *Psychol. Rev.*, 60, 1953, 33-39.

single-trial versus incremental learning interested him, particularly after his publication in the *JOURNAL* of Rock's study in 1957.¹⁹ His final experimental article dealt with a stochastic model for recall in a verbal-learning paradigm.²⁰ He was still active and alert, but the years began to take their toll as he neared 80.

In the mid 1960s the University of Texas passed a new regulation requiring retirement at the age of 70, but an exception was made for Dallenbach, and he continued to teach until ill health forced him to retire at 80.

Dallenbach's tenacity was probably his greatest characteristic, and perhaps also his weakness. Once he set his mind toward a goal there was no diverting him. It made giving up his responsibilities at the end of his professional career all the more difficult. This was particularly true with his giving up the *AMERICAN JOURNAL OF PSYCHOLOGY*. He had talked about divesting himself of it since the beginning of the 1960s. He received several offers to buy it, one a six-figure offer from a major publisher. He knew it had to be done if only to guarantee the future of the *JOURNAL*, but he kept putting off the decision as long as he could. "I can't let go! Why can't I let go?" he exclaimed once in the late 1960s. It was simply difficult to give up a life's work, but particularly difficult for a man who did not believe in giving up.

The *JOURNAL* was largely responsible for Dallenbach's research efforts, as large as they were, not being larger. The drain of it was also largely responsible for the fact that the book on attention was never written, nor the one on psychophysics (which he suggested J. P. Guilford write instead). He sometimes regretted that he never wrote those books, but he would wave his hand at the *JOURNALS* that lined his shelves and say, "Those are my books." It was his plan after retirement to gather together the necrologies in this *AMERICAN JOURNAL* series and publish them as a separate book. Ill health prevented him from completing the task, but hopefully, his plan will be realized in the near future.

At long last, as his health began to fail him, he realized that something had to be done with the *JOURNAL*. But where was it to go? Its traditions were very important to him and he had always lived up to the ideals Titchener had laid down—that it not be operated for profit, that what profits resulted were to be pumped back into it to improve it,

¹⁹I. Rock, The role of repetition in associative learning, this *JOURNAL*, 70, 1957, 186–193.

²⁰R. B. Evans and Karl M. Dallenbach, Single-trial learning: A stochastic model for the recall of individual words, this *JOURNAL*, 78, 1963, 545–556.

and that it be operated for the good of experimental psychology. He always operated with more interest in its quality than in the cost. He used the largest number of type faces of any psychological journal, doing this to preserve the carefulness of language that was the hallmark of Dallenbach's JOURNAL.

He decided that the JOURNAL had to be in the hands of a group that would not operate it for profit; one that would be willing to maintain its quality production, even if it meant just 'breaking even' financially; and one that would attempt to maintain its traditions even if they were not always journalistically expedient. He considered the Psychonomic Society and the University of Texas, but he finally decided on his Alma Mater, the University of Illinois. In 1968 the gift was made, relieving Dallenbach after 48 years of his JOURNAL responsibilities.

The Dallenbach of Texas was different from the one of Cornell. He was mellow and perhaps a bit less aggressive. Students at first exposure were somewhat awed by his Teutonic figure and his formality.²¹ It did not take long, however, for the discerning student to discover that the formality was his nineteenth-century etiquette and that behind it was a kindly, friendly person. He seemed to adopt his doctoral students, worrying after them as though they were his own children. He was always ready to 'go to bat' for one of his students. One now-famous physiological psychologist, a student of Dallenbach's at Cornell in the 1930s, recalled how, after he had 'flunked out' of Cornell, Dallenbach managed to get him reinstated and planned out with him ways to develop study habits.

At Texas Dallenbach was referred to as 'Papa Bear'—behind his back, of course. It indicated how aware students were of the thinness of his external formality. In his later years, he came to the campus only a few days a week, preferring to do most of his editorial work at home, where he had access to his excellent library of psychological periodicals. The graduate student, often as not, saw him at home. It was all in the image of Titchener, of course, but the reasons were practical, not sentimental.

In the Texas years, the highlight for many a student of Dallenbach's history or psychophysics courses was the end-of-semester social hour. Dr. and Mrs. Dall would always have the class over one evening to

²¹Once, when Margaret McGrade was introduced to a University of Texas student as Dr. Dallenbach's secretary, he responded, "Oh, then you must speak German very well." When asked why this must be so, the student replied, "Well, everyone knows that Dr. Dallenbach is a German and speaks nothing but German."

"get to know the students better." Dr. Dall would often show his home movies of the Experimentalists or the films of the facial-vision experiments. There was always good food—often delicacies sent to the Dallenbachs by ex-students—and pleasant conversation. In a large, impersonal university, it was one of the rare moments when an undergraduate could get to know one of the 'big men' in the department.

One of the things Dallenbach always wanted to do was to lecture to his class in his academic robes. He managed to do it only vicariously, by having this biographer promise to offer a history of psychology course at least once in academic robes. The promise was made, and it has been followed ever since.

His honors were many but he always played them down. Awards embarrassed him somewhat, but the honors that probably meant the most to him were the Gold Medal of the APA, the Dallenbach Commemorative Number of the AMERICAN JOURNAL OF PSYCHOLOGY in 1958, and his honorary doctorate from the University of Illinois in 1968.

Dallenbach's view of religion since his days as a graduate student was that of a skeptic, albeit a good natured one. His views were influenced strongly by his reading of Winwood Reade's *The Martyrdom of Man*. The primary reason for his view, however, was linked to his own character. He believed that religion teaches a person to do what is right in order to gain a reward in another life. "A person shouldn't do what is right just to get something out of it," he said once. "A person should do what is right because it is the right thing to do." That was Dallenbach's creed and better than anything else typifies his life.

Rand B. Evans, *University of New Hampshire*

SCALAR PERCEPTIONS WITH BINOCULAR CUES OF DISTANCE

Walter C. Gogel

University of California, Santa Barbara

Under reduced conditions of observation, the perception of egocentric distance is determined by a composite of the tendency to see objects as near (the specific-distance tendency) and residual oculomotor cues of distance. The resulting perceived distance—the egocentric reference distance—was found to affect the perceived relative depth produced by binocular disparity between two points of light: the more distant light was positioned in apparent depth near the egocentric reference distance, with the scalar perception of the depth between the lights related to the magnitude of the reference distance. Implications for the scaling of relational perceptions are discussed.

Scalar perceptions of extent (size or distance) are perceptions that are metric in character. Examples are the perceptions that an object is 3 feet from the observer, is 3 inches tall, or is 2 feet closer than a second object—or, in more direct, behavioral units, that an object is at arm's length or is small enough to be held in the hand. These perceptions are contrasted with nonscalar perceptions, such as that one object is twice as distant as another, without regard to the magnitude of this distance, or is three times the size of another object, without concern as to the size of either object (Gogel, 1968; Ono, 1970).

Most research in spatial perception has involved nonscalar perceptions. This includes all research in which a variable and a standard object are simultaneously or successively equated in size or distance to be some perceived multiple of each other, as in the research on size constancy, where usually a variable at a constant distance is perceptually equated in size to a constant standard at a variable distance. The results of such size-constancy studies, although indicating whether perceived size changes with physical distance, give no indication of perceived metric sizes.

It has been pointed out that conclusions from studies of perceived size can differ depending on whether scalar or nonscalar measures are obtained (Gogel, 1968). Also, the lack of research on scalar perceptions should not be taken as an indication that they are unimportant. It is obvious that scalar perceptions are essential to the ability to respond

meaningfully in many situations. Here is one instance.

Consider the situation in which a patterned surface (floor) of 2-in. checkerboard squares is perceived as extending in depth, with its nearest edge at some distance from the observer. Assume that the task is to report scalar sizes and distances within this display. Clearly, without some metric information or some metric assumptions this task cannot be performed.

It is possible that if an off-sized familiar object, such as a very small chair, were introduced into the display, the observer might perceive the checkerboard squares as the size of floor tiles, perhaps as 9-in. squares. On the other hand, if instead of an off-sized chair, a normal-sized chess figure were the only familiar object present, the observer might perceive the squares as about 2 in. on a side. In these cases, the familiar object might be said to provide a metric, or ruler, for scalar perceptions throughout the visual field.

Within a checkerboard pattern, there are cues of relative size and relative distance (a gradient of retinal size) to determine nonscalar perceived depth and nonscalar perceived size. The introduction of metric information anywhere within the configuration may possibly trigger scalar perceptions throughout the configuration, with the veridicality of the scalar perceptions depending on the veridicality of the metric, or ruler, as the chair/chessman example suggests. Although the author is not aware of any studies demonstrating such a triggering process, the example serves to raise the issue of the process by which relational information is transformed to scalar perceptions.

There is evidence to suggest that a familiar object would *not* be very effective for such a triggering, since it is doubtful that familiar size is a good determiner of perceived size, either in visually rich environments (Fillenbaum, Schiffman, and Butcher, 1965; Schiffman, 1967) or in visually reduced environments (Gogel, 1969a; Gogel and Newton, 1969). But when cues of relative distance are available, even under conditions in which all familiar-size cues that might provide scalar information are absent, the observer can nevertheless make scalar responses.

This latter effect can be demonstrated with the relative-size cue to distance occurring between two monocularly observed objects of the same shape, but different visual angle, that are presented either simultaneously or successively. In this case, scalar reports of the distance between the two objects can be made—reports which, although not necessarily veridical, vary systematically with the ratio of the visual angles of the objects (Epstein and Landauer, 1969; Gogel, 1969c). Or if binocular disparity is present between two points of light in an other-

wise dark room, as the present study will demonstrate, scalar judgments of the depth between the points of light can occur.

Indeed, it might be generalized that scalar perceptions *always* occur to depth intervals generated by relational cues. In other words, an observer may never perceive a nonscalar depth (or size) even though obvious sources of scalar information seem to be lacking. The resulting scalar perception of depth might not be proportional to the physical distance between the objects but neither is it randomly determined. A possible source of scalar information under such conditions may lie in the oculomotor cues of the accommodation and convergence of the eyes.

Oculomotor cues to distance

In considering accommodation and convergence as a possible source of scalar information, two kinds of accommodation and convergence should be distinguished. One of these, absolute accommodation and absolute convergence, results from the eye adjustments to the particular distance of an object. The other, relative accommodation and relative convergence, consists of accommodation and convergence changes between objects presented either simultaneously or successively (see Gogel, 1961a; Ono, 1970).

It is possible, for example, that relative accommodation or convergence can indicate that one object is more distant than another even under conditions in which absolute accommodation or convergence is an inadequate cue to the distance of the object from the observer. But a given accommodative or convergence difference is not uniquely related to a specific depth interval: an accommodative difference of $\frac{1}{2}$ diopter can result from a distance change of 40 to 50 cm or from a distance change of 100 to 200 cm, for example. Therefore, it is unlikely that an accommodative or convergence *difference* can provide scalar information.

However, since a unique value of *absolute* convergence or accommodation is associated with a particular distance, absolute accommodation or convergence can possibly provide scalar information as to the egocentric distance of an object. And since it is usually found that convergence is a more effective cue than accommodation (see Heinemann, Tulving, and Nachmias, 1959), absolute convergence is more likely than absolute accommodation to provide the scalar egocentric-distance information that would serve as the metric, or ruler, for the scalar calibration of depth intervals that are produced solely by nonscalar, relational cues.

If absolute convergence or accommodation were to determine the

perceived scalar distance to *each* end of a depth interval (such as a checkerboard floor) though, the perceived interval would not be produced by nonscalar, relational cues. In this case, the scalar perception of the depth interval would present no explanatory problem, since it would result from the difference between two scalar distance perceptions. Thus, to investigate the calibration of a perceived depth interval produced by nonscalar, purely relational cues, that interval must be supported by purely relational cues—an effective scalar egocentric-distance cue must not be available to both ends of the display.

Specific-distance tendency

Absolute convergence as a source of scalar information is limited in that this cue is usually considered to be ineffective for far distances. It has often been found that the distances of objects about 10 ft and farther from the observer are underestimated when only the cue of absolute convergence is available. For example, in a study by Foley (1968), which used convergence values from divergence to approximately 1 m, the reported distance of an object varied from 5.2 to 1.6 ft for one observer and from 4.2 to 1.9 ft for another. In a study of whiteness contrast between a binocularly viewed large and small disc, with the large disc at 9.3 ft from the observer, Gogel and Mershon (1969) found that the average reported distance of the large disc was 3.8 ft. In a study of perceived size as a function of convergence and apparent distance, Biersdorf (1966) found that a binocularly presented object at a distance of 4 m was reported to be at distances of about 21, 20, 9, and 8 in. by the four observers.

This underestimation of distant objects with convergence is part of a more general process termed the specific-distance tendency—the tendency under reduced conditions of observation (that is, in the absence of cues of egocentric distance) for objects to appear at a relatively near distance (about 6 ft) from the observer, regardless of their physical distances. If the visual conditions are only partly reduced, the cues that remain (oculomotor cues of distance, for example) will somewhat modify the perceived distance expected from the specific-distance tendency (Gogel, 1969c).

There are several sources of evidence for the specific-distance tendency. As just indicated, under reduced or fairly reduced conditions of observation distant objects tend to appear nearer than their physical distances. Also, if an object is presented with distance information so as to appear at a far distance and the distance information is later removed, the object decreases in both perceived egocentric distance

and perceived size (Gogel, 1969c). Furthermore, if familiar objects are presented at various simulated distances, the objects that are at simulated far distances often appear to the observers to be smaller than normal in size and the objects that are at simulated near distances often appear to be larger than normal in size (Gogel, 1969b; Gogel and Newton, 1969). That is, the result of the specific-distance tendency is to make these near and far familiar objects appear at a similar, relatively near distance, with the perception of their sizes determined by this apparent distance.

This result with familiar objects is supported by an experiment by Gogel and Mertens (1967), in which the observers adjusted a playing card to 3, 6, or 9 ft. There were no distance cues available except familiar size. The stimulus was an image of a playing card at a constant distance of 400 cm, whose size could be controlled by the observer. The starting size (starting visual angle) for the adjustment was determined by the experimenter. It was found that the adjusted (simulated) distances, as computed from the adjusted sizes, varied inversely with the starting size. It was as though the perceived distance associated with the starting size was displaced from the simulated distance to a distance of about 7 ft from the observer. As a result of this initial error in perceived distance, the simulated distance required to achieve a perceived distance of 3, 6, or 9 ft varied systematically as a function of the starting size.

It can be concluded that even in the absence of *conventional* cues to egocentric distance, an egocentric-distance perception will occur as defined by the specific-distance tendency. Thus, the specific-distance tendency is also (like oculomotor cues) a possible source of information that would serve as the metric, or ruler, for the calibration of depth intervals produced by relational cues.

Egocentric reference distance

Although the specific-distance tendency can be considered conceptually as independent of oculomotor cues of distance, in any situation it is likely that some residual oculomotor cues will be present. It is asserted here that the specific-distance tendency together with residual cues from convergence and accommodation provide a scalar perception of egocentric distance that acts as a metric, or ruler, permitting relational information to be converted to scalar perceptions. This scalar perception of egocentric distance that is determined by the composite effect of the specific-distance tendency and oculomotor cues is here termed the egocentric reference distance. Only in the complete absence

of oculomotor cues to distance is it expected that the egocentric reference distance will be determined solely by the specific-distance tendency.

To understand the process by which scalar information is introduced into a relational perception of depth by the egocentric reference distance, it is necessary to understand how a perceived *depth* interval is positioned with respect to the perceived *distance* defined by this reference.

Consider again the example of the perceived depth interval produced by a checkerboard floor when all the usual cues to the egocentric distance of that floor have been removed. Suppose that the ratio of the retinal sizes of the nearest to the farthest edge of the floor is 2 to 1. If the most distant portion of the floor appears at the egocentric reference distance of 6 ft, the nearest portion will appear at half that distance and the perceived depth along the floor will be 3 ft. On the other hand, if the *nearest* portion of the floor appears at the egocentric reference distance, the farthest portion will appear at twice this distance (12 ft), with the perceived depth along the floor equal to 6 ft. Since all portions of the floor cannot appear simultaneously at the reference distance, the apparent position of the checkerboard floor with respect to that reference distance is a significant variable in determining the scalar perception of the depth along the floor. The main objective of the present study was to investigate how a relationally produced depth interval within a configuration is perceptually positioned with respect to the egocentric reference distance.

In this study, the perceived depth interval, instead of being generated by relative-size cues in a checkerboard floor, was produced by binocular disparity between two points of light presented simultaneously at different stereoscopic distances. Binocular disparity, since it consists only of a difference in horizontal separation in the two eyes, by itself provides no scalar information (Gogel, 1969c; Ono, 1970).

As noted above, it was essential for the study that no scalar cues be available to modify directly the perceived distance between the two ends of the depth interval—in this case, between the two (simulated near and far) lights. For example, if the cue of absolute accommodation or convergence were available to determine an apparent scalar distance to *each* of the lights, the occurrence of a scalar perception between them would be expected and would tell us nothing about the scalar calibration, by way of the egocentric reference distance, of a relationally produced depth interval. Therefore in the present study the *accommodation* to the two lights was constant.

It was also necessary to avoid the possibility that absolute *convergence* to each of the lights (rather than binocular disparity) produced the perceived depth between them. This requirement was examined by presenting each (near or far) light singly to two separate groups of observers, as well as presenting both (near and far) lights simultaneously to a third group of observers. If the average reported distances of the lights presented singly were not different between the first two groups, it would indicate that the perceived depth obtained from the group presented the two lights simultaneously was not itself produced directly by the absolute-convergence cue to the scalar egocentric distance of both lights. In addition, under these circumstances the average reported distance of the singly presented lights would provide a measure of the egocentric reference distance, that is, a measure of the specific-distance tendency modified somewhat by the average value of convergence.

EXPERIMENT I

Method

—Apparatus—A Polaroid stereoscope was used to present points of light (1.0 mm in diameter) at simulated distances of either 5 or 25 ft from the

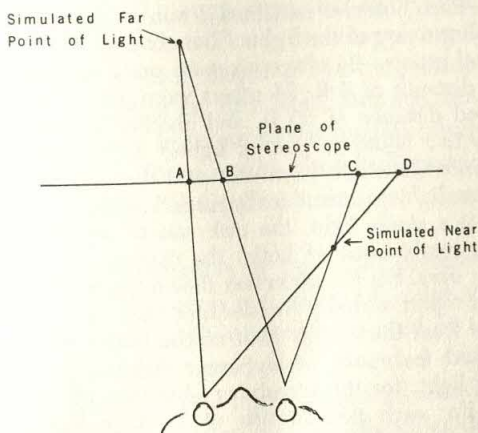


Fig. 1. Schematic diagram of the Polaroid stereoscope used to produce the simulated distances of the two points of light; the binocularly viewing observer is represented at the bottom of the diagram

observer. The brightness of all the lights was 1.25 log units above the absolute threshold (as determined after a 2-min period of dark adaptation). The physical position of the lights (the accommodative distance) in Experiment I was constant at 8.33 ft from the observer, with the lights horizontally separated by 7 deg 28 min of visual angle. The apparatus is illustrated schematically in Figure 1. Appropriately oriented Polaroid filters were located both at the plane of the stereoscope and immediately in front of the observer's eyes. In the situation illustrated in Figure 1, the Polaroids were oriented so that points A and D stimulated the observer's left eye and points B and C his right eye, with the result that the left point was at a greater simulated distance (25 ft) than the right point (5 ft). Reversing the orientation of the Polaroids in front of the eyes resulted in a simulation in which the right object was at 25 ft and the left at 5 ft.

The two lights could be presented *simultaneously*, or either the left or right light could be presented alone (*singly*). Care was taken that nothing was visible to the observer except the lights, with the remainder of the visual field kept totally dark. The observations were made from a dark booth containing a head-and-chin rest, the Polaroids, and a shutter for occluding the stimuli. A microphone and earphones permitted the experimenter and the observer to speak to each other.

—Observers—The observers were 72 students in an introductory course in psychology. Each had an acuity (corrected, if necessary) of at least 20/20 in each eye (near and far) and in addition a macular stereoscopic acuity of at least 22.08 sec of arc as measured with a Keystone stereoscope. All were naive as to the purpose of the experiment.

—Procedure—Each observer remained 2 min in the dark observation booth before being shown any of the lights. The observers were randomly divided into three equal groups: 24 observers were presented only a single light at the simulated distance of 5 ft, 24 others were presented only a single light at the simulated distance of 25 ft, and another 24 were presented simultaneously with two lights (a pair of lights) at the simulated distances of 5 and 25 ft. Binocular observation was used throughout, with the right/left positions of the lights systematically varied within each group. For the observers shown a single light, the task was to report verbally (in feet or inches or some combination of both) the distance that the light appeared to be from their eyes. For the observers shown the two lights simultaneously, the task was to report verbally the distance that each of the points of light appeared to be from their eyes. Half of the observers presented a pair of lights were asked to report the apparent distance of first, the near, and second, the far light; for the remaining observers of this group, this order was reversed. For each presentation of a single light, the experimenter ascertained, by questioning, that the proper binocular fusion had occurred before the observer gave a report of apparent distance. When presented with a pair of lights, the observer was asked to look from one light to the other and to indicate which appeared more distant before reporting the apparent distance of either light.

Results

The results of Experiment I are shown in Table I. Medians rather than means were considered to best represent the data, since some of the distributions are skewed. This is particularly clear for the near presentation of the single light, since in this case when the results from three observers are removed, the mean and standard deviation become 4.40 and 3.06 respectively.

Table I. Means, medians, and standard deviations, in feet, of the reported distances of the points of light for the single and simultaneous presentations at the near (5 ft) and the far (25 ft) simulated distances, Experiment I

	Single light		Two lights simultaneously	
	Near	Far	Near	Far
Mean	83.02	8.78	3.33	7.88
Median	4.00	7.50	3.00	7.50
SD	301.74	6.17	2.10	4.66

An overall result evident in Table 1, and expected from the *specific-distance tendency*, is that the median results underestimate the simulated egocentric distances of the lights, particularly in the case of the 25-ft simulation. Verbal underestimations of egocentric distance frequently occur even in multiple-cue conditions of observation (Gogel, Hartman, and Harker, 1957), suggesting that some component of the underestimation can be attributed to errors in applying the memory of a foot ruler to the perceptions of distance. The magnitude of the underestimation of the light at the far simulated distance, however, is such as to make any explanation in terms of the mental misapplication of a foot ruler highly unlikely. These results, therefore, provide evidence of the effectiveness of the specific-distance tendency under the conditions of this experiment.

If the median reported distances of the lights presented singly (the single lights) had been the same regardless of convergence distance, that is, regardless of whether they were presented at the near or far simulated distance, then this would have provided evidence that the perceived depth between the two lights presented simultaneously (the simultaneous lights) was supported only by binocular disparity. If this had happened, the median response to either of the single lights would have provided a measure of the *egocentric reference distance*, a value that could then have been applied to the results from the simultaneous lights in order to determine which of them appeared at or near the

reference distance. However, a test of the medians for the single lights indicated (using the Mann-Whitney U test) that the bulk of the reported distances of the far lights was greater than that of the near lights ($Z = 2.23$, $p < .02$). Therefore, the question of how the two simultaneous lights were positioned perceptually with regard to the egocentric reference distance could not be answered by Experiment I.

In the case of the two *simultaneous* lights, it may be noted, the reported distance of the far light was greater than the reported distance of the near light, for every presentation of the pair of lights. It should also be noted that the median perceived distance of the far light was the same whether it was presented by itself or simultaneously with the near light but that this was *not* the result for the light at the near simulated distance. The bulk of the perceived distances of the near light was significantly greater (as determined by the Mann-Whitney U test, $Z = 2.26$, $p < .02$) when the near light was presented singly than when it was presented together with the far light. If the median value of the egocentric reference distance here is assumed to be 7.5 ft, these results suggest that the far light was always perceived to be near that reference distance (and thus unchanged in apparent distance whether presented singly or simultaneously with the near light) but that the binocular disparity between the simultaneous lights caused the near light to appear closer than the constant perceived distance of the far light (and thus closer when presented simultaneously than singly).

To investigate the scalar calibration of a relationally determined depth interval it is required that the perceived depth interval is indeed relationally determined—that it was not produced or modified by scalar cues such as absolute convergence or accommodation. Since the difference in the reported distances of the single (near versus far) lights for the first two (near versus far presentation) groups of observers can account for much of the perceived depth between the simultaneous lights by the third group, this requirement was not met. The results are nevertheless useful in that they demonstrate that absolute convergence *can* be a cue to egocentric distance, at least for distances as different as 5 and 25 ft.

The problem of whether convergence is a cue to distance has generated a long history of often conflicting results (see Biersdorf, 1966; Biersdorf, Ohwaki and Kozel, 1963; Gogel, 1961a, 1961b, 1962; Heine-mann, Tulving, and Nachmias, 1959; Richards and Miller, 1969; Cran-nell and Peters, 1970). In the present study the use of the two groups of observers each with a particular convergence distance assured that the cue of absolute convergence—as distinct from the cue of conver-

gence differences—was being investigated. Although it is clear that considerable variability occurred between observers, the results support the ability of this cue to discriminate between egocentric distances of 5 and 25 ft.

EXPERIMENT II

Only if absolute convergence had been ineffective in Experiment I in discriminating between the two simulated distances of the singly presented points of light—only then could the average values of these perceived distances have been used to determine how a perceived depth interval produced by binocular disparity is positioned with respect to the egocentric reference distance. Experiment II attempted to achieve this condition by reducing the difference between convergence distances for the near versus far lights. It seemed reasonable to expect that despite this reduction, the binocular disparity in the simultaneous presentations would still be an effective cue that the 'farther' light was at a greater (simulated) distance than the 'nearer' light.

Method

—Apparatus—In Experiment II, the simulated distances of the 1.0-mm points of light were 10 and 20 ft, rather than the 5 and 25 ft of Experiment I. The brightness of the lights was the same as in Experiment I. The physical (accommodative) position of the lights (the plane of the stereoscope) was constant at 13.33 ft from the observer, with the lights horizontally separated by 2 deg 20 min of visual angle. In all other respects, the apparatus was identical to that of Experiment I. Also, in order to obtain a measure of the specific-distance tendency unaffected by oculomotor cues, a point of light at a physical distance of 13.33 ft was presented singly to some observers and was viewed *monocularly* through a 1.0-mm restrictive pupil (with the Polaroids absent). The luminance of this light viewed through the artificial pupil was adjusted until it appeared to the experimenter to have the same brightness as the lights viewed through the Polaroids.

—Observers—The observers were 120 students in an introductory course in psychology. They met the same acuity requirements as in Experiment I. None had been used in Experiment I and all were naive as to the purpose of the experiment.

—Procedure—The task and instructions were the same as before. Five groups with 24 different observers in each group were used, with observers randomly assigned to the groups. One group was first presented the single light at the near (10 ft) simulated distance, followed by the simultaneous lights at the simulated distances of 10 and 20 ft. Another group was presented first the single light at the far (20 ft) simulated distance, followed by the simultaneous presentation of the pair of lights at the simulated

distances of 10 and 20 ft. Another group was presented first the pair of simultaneous lights and then the single light at the near (10 ft) simulated distance. Another group was presented first the pair of simultaneous lights, then the single light at the far (20 ft) simulated distance. Another group was presented only the single light, viewed monocularly through the artificial pupil. The order in which the perceived distances of the near and far lights in a pair were reported and the order in which the near or far light was on the right or left in either the single or simultaneous presentations were counterbalanced within groups.

In summary, Experiment II was similar to Experiment I, the main differences being that the simulated distance (convergence difference) between the points of lights was decreased, that second as well as first presentations were used, and that a separate group of observers monocularly viewed a single light through a restrictive aperture. Except for this last group, all viewing was binocular.

Results

The results of Experiment II for all groups except the group using the artificial pupil are shown in Table 2. The mean, median, and standard deviation of the results with the artificial pupil were 4.60, 3.00, and 4.84 ft respectively.

Table 2. Means, medians, and standard deviations, in feet, of the reported distances of the points of light for the single and simultaneous presentations at the near (10 ft) and far (20 ft) simulated distances, Experiment II

	Single light		Two lights simultaneously			
			Near reported first		Far reported first	
	Near	Far	Near	Far	Near	Far
First presentations						
Mean	7.54	37.25	4.73	7.83	4.38	6.54
Median	5.50	6.00	4.00	6.50	4.00	6.50
SD	4.87	82.46	3.55	5.83	1.47	2.26
Second presentations						
Mean	7.48	6.88	6.61	11.44	12.40	19.38
Median	7.50	5.50	4.50	7.00	4.50	7.00
SD	3.47	4.93	6.04	11.60	39.15	58.60

As in Experiment I, when the two lights were presented simultaneously the far light was perceived to be more distant than the near light by all observers. However, unlike Experiment I, the data for the first presentations show that when the lights were presented singly the bulk of the perceived distances of the far light was *not* significantly greater than that of the near light (Mann-Whitney U test, $Z = .12$, $p > .05$). From these results and the median values of Table 2, it can

be concluded that the difference in convergence to the single (near versus far) lights was below the threshold for discrimination on the first presentations, whereas on the other hand, and as expected, the binocular-disparity cue between the simultaneous lights consistently resulted in a difference in perceived distance on both the first and the second presentations. It will also be noted that the means and medians for the single lights on the second presentations were greater for the near than for the far light—though this difference, which is opposite in direction to that expected from the convergence cue, was not significant at the .05 level ($Z = 1.02$, using the Mann-Whitney U test).

As in Experiment I, the large median underestimations of the far simulated distance (with either singly or simultaneously presented lights) provides evidence for the effect of the *specific-distance tendency* on the perception of distance. But, unlike Experiment I, the reported distances of the single lights at the two simulated distances (near versus far) were *not* significantly different. It follows that the average median value (5.75 ft) from the first presentations of the single lights is an estimate of the *egocentric reference distance* and can be applied to the perceived depth resulting from binocular disparity.

This 5.75-ft average is between the median values of 4.0 and 6.5 ft obtained from the *simultaneous* lights and is somewhat closer to the far than to the near light of this pair. Further, on these first presentations, the bulk of the perceived distances of the near light singly presented was again, as in Experiment I, significantly greater (Mann-Whitney U test) than the bulk of the perceived distances of the near light simultaneously presented ($Z = 3.26$, $p < .01$), but the bulk of the perceived distances of the far light singly presented was again, as in Experiment I, *not* significantly less than that of the far light simultaneously presented ($Z = .07$, $p > .05$). Thus it seems that a pair of simultaneously presented lights is positioned in apparent depth such that the egocentric reference distance is closest to but somewhat less than the farther light in the pair.

This conclusion is supported by the results from the second presentations. In this latter case, the measure of the egocentric reference distance (an average median of 6.50 ft) was less than but close to the median perceived distance (7.00 ft) of the far light, and considerably farther than the median perceived distance (4.50 ft) of the near light of the simultaneously presented pair. In other words, according to the results from Experiment II, under conditions of reduced egocentric cues to distance, a depth interval assumes an apparent position with respect to the egocentric reference distance such that this reference

distance is less than but near to the apparent distance of the far light.

The median reported distance (3 ft) of the light viewed *monocularly* through the artificial pupils might also be considered a measure of the specific-distance tendency, in this case unbiased by oculomotor cues. It has been found, however, that when the stimulus to accommodation is removed by presenting a homogeneous visual field, the accommodative state of the eye is to a relatively near distance (Schober, 1954). It is likely, therefore, that the use of the restrictive pupil, by removing the need for the eye to accommodate, resulted in an accommodation (and an accommodative convergence) to a near distance. On the other hand, it is possible also that both the resting state of accommodation and the unbiased value of the specific-distance tendency are to the same distance, indicating perhaps the special behavioral significance of this near distance (see Schober, 1954).

DISCUSSION

Calibration of the depth interval from binocular disparity

The main conclusion from this study is that a depth interval generated solely by binocular disparity between two lights is so perceived that the farther of the two lights appears near the egocentric reference distance. As a result of this process, the observer is able to perceive the depth interval as a scalar distance. Basic to this conclusion is the assumption that the reports of scalar distance obtained under the restricted conditions of observation of this study were meaningful.

If *meaningful scalar perceptions* occurred between the simultaneously presented lights, these perceptions should have varied as a function of the amount of binocular disparity and the several different reports of distance for the same observer should be correlated. Comparing the results from Experiments I and II, it is clear that when the binocular disparity between simultaneously presented lights was greater, the perceived depth between the lights was also greater. The consistency between the different reports of distance was examined by calculating Pearson product-moment correlations between the perceived distance of the far point of light and the perceived depth between the lights, for each observer shown the simultaneous presentations. (The perceived depth between the lights was obtained by subtracting the perceived distance to the near light from the perceived distance to the far light.) The resulting values of r were .93 ($N = 24$) for Experiment I and .86 and .99 ($N = 48$) for the first and second presentations of

Experiment II. Similar correlations obtained for the perceived depth between the lights and the perceived distance to the *near* light were .57 and .58 and .98 respectively. All of these correlations were statistically significant at the .01 level. These correlations, particularly those involving the far light, provide clear evidence that the scale associated with one part of the configuration of light was applied consistently throughout the configuration.

In addition, if the *egocentric reference distance* (as measured by the perceived distance of single lights in Experiment II) did calibrate the perceived scalar depth between pairs of simultaneous lights, these two kinds of perceptions should be correlated positively. In other words, the distance at which an observer perceived the single light should be correlated with the depth he perceived between the pair of simultaneous lights. Since the former measure was skewed, the Spearman rank-correlation coefficient corrected for ties was applied to the results from all 96 observers regardless of whether the single light or a pair of lights was presented first. The resulting correlation was +.68, which was significant beyond the .01 level.

The results from the present study are consistent with the following explanation of the scalar perception of a depth interval. When an observer is presented a configuration of binocularly observed objects distributed in apparent depth, the stereoscopically more distant part of this configuration is placed perceptually near the egocentric reference distance (about 6 ft in the present study), with the remainder of the configuration appearing to extend toward him from this perceived distance. Thus, the perceived scalar depth from a binocular disparity is a function not only of the magnitude of the binocular disparity but also of the magnitude of the egocentric reference distance. As is indicated by the correlation of +.68, the greater the magnitude of the egocentric reference distance (for a constant binocular disparity), the greater will be the perceived scalar depth within the configuration.

The results from the first presentations of the simultaneous lights can also be applied to test specific hypotheses of binocular vision. According to Von Kries (1925),

$$\frac{id'}{D'_n D'_f} = \Gamma, \quad [1]$$

where Γ is the binocular disparity, d' is the perceived depth between the two points of light, D'_n and D'_f are the perceived distances of the near and far points of light from the observer, and i is the interpupillary distance (which was assumed in the calculations that follow to be a con-

stant). According to Equation 1, the ratio of the left-hand quantities obtained from Experiments I and II should be equal to the ratio of the binocular disparities in these two experiments. The ratio of the disparities was 3.2, whereas the ratio of the left-hand quantities was 2.2 from the means and 2.1 from the medians.

A hypothesis consistent with the Luneburg theory of binocular vision (see Foley, 1967) is that

$$d'/D_n = K\Gamma, \quad [2]$$

where K is an observer constant. The ratio of the left-hand quantities of Equation 2 using either means or medians from the present two experiments was 2.4, again as compared with the ratio of the binocular disparities of 3.2.

From an experiment in which estimates were obtained of the ratio of the perceived distances of a near and a far light, Foley (1968) has suggested modifying Equation 2 by raising Γ to a power P . But the value of P calculated from the present study was less than unity, whereas Foley, using three observers, has found values of P greater than unity. Possibly the use of naive observers making only one or several distance responses, as in the present study, does not produce results that are equivalent to the results obtained from a more extensive series of judgments with the same observers.

Or possibly the d' between the two simultaneous lights in Experiment I was determined not only by binocular disparity but also by the absolute-convergence cue to the distance of each light. It has been found that the function relating perceived distance and convergence can be different for different objects in a binocular configuration (Gogel, 1961b). It follows that an observer who is binocularly presented two lights, with at least one light at a relatively near distance, has two discrepant sources of information for the perception of d' ; one from convergence and the other from binocular disparity. If both sources contribute to the perception of d' , the value of d' could possibly differ from that which would occur if only binocular disparity were available to determine d' .

Finally, the results from a previous experiment by Gogel (1961b, Experiment II) are consistent with and suggest an extension of the present results on the apparent positioning of a binocular configuration with respect to the egocentric reference distance. In that experiment, the relation between convergence and perceived distance was investigated using a monocularly observed comparison field and either a single object or a configuration of three simultaneously presented binoc-

ular objects. It was found that the function relating convergence and perceived distance for the single object was more like that for the middle and far than for the near object of the configuration. Although convergence was somewhat effective with half of the observers and, although the egocentric reference distance therefore changed with changes in convergence, the results were that the farther portion of the binocular configuration tended to appear at the distance of this somewhat variable egocentric reference distance. This suggests that the results obtained in Experiment II of the present study also would have occurred if three rather than two points of light had been simultaneously presented; that is, that the far light in this case also would have appeared near the egocentric reference distance.

Application to other relative-distance cues

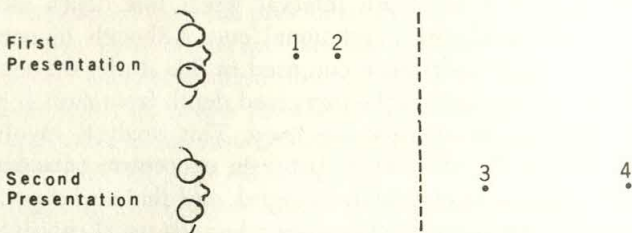
This paper has examined the process by which a metric is introduced into a perceived depth interval when this depth interval is determined exclusively by a relational cue. Although binocular disparity was the relative-distance cue used in this study, the same analysis could also be applied to the perceived depth from motion-parallax, relative-brightness, or relative-size cues. This analysis involves two assumptions: that the distinction between egocentric (absolute) and exocentric (relational) cues is meaningful, and that scalar perceptions of extent occur to exocentric cues even when the usual egocentric cues are absent. Although the distinction between egocentric and exocentric cues has not been widely investigated, there is some evidence that it is valid in a number of situations (Gogel, 1961a, 1969a; Kinchla, 1971). The second assumption is supported in that scalar perceptions do occur under highly reduced conditions of observation. Although these perceptions often are nonveridical, they vary systematically as a function of stimulus parameters, or of the specific-distance tendency if none of the usual egocentric cues are present.

One consequence of those assumptions is that scalar perceptions are involved in adjusting different depth intervals to perceptual equality or to a perceived ratio. If the perception of depth intervals is necessarily scalar even under reduced conditions of observation, the depth intervals will be reported as equal, for example, only when they are perceived as having equal scalar extents. According to Equations 1 and 2, perceived depth from binocular disparity is calibrated by the perceived distance of the configuration.¹ According to the results of the present study, it is calibrated by the egocentric reference distance in the ab-

sence of the usual egocentric cues. Also, the concept of an egocentric reference distance is not limited to binocular observation and could be important in the perceived depth associated with any relational cue.

The hypothesized interrelation between the egocentric reference distance and the scalar perception of a relationally produced depth interval is illustrated in Figure 2. Figure 2 is intended to represent a situation in which a relational cue—in this case, a constant binocular disparity—is present between objects 1 and 2 and between objects 3 and 4, with none of the usual egocentric cues to the distance of either of these configurations. As indicated in Figure 2, the configurations are presented successively in order to avoid binocular disparity between the two configurations per se. Since the egocentric reference distance in this case is assumed to be determined entirely by the specific-distance

PHYSICAL



PERCEIVED

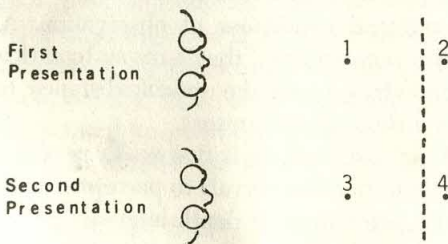


Fig. 2. Physical and apparent depth between two points of light observed binocularly for the condition in which the egocentric reference distance is assumed to be the same as the perceived distance determined by the specific-distance tendency; the observer is represented at the left, the specific-distance tendency by the broken line

tendency and since the binocular disparity between the two pairs of objects is the same, the perceived scalar depth as illustrated in Figure 2 is the same for the two configurations.

If the egocentric reference distance were greater or less than the specific-distance tendency (than the broken line, Figure 2), then the perceived depth between the pairs of objects would increase or decrease respectively. Or if the magnitude of the binocular disparity between one pair of objects (for example, between objects 1 and 2) were greater than that between the other pair of objects (3 and 4), the perceived scalar depth between the former would be greater than that between the latter. Or unequal binocular disparities could result in equal perceived distances if the egocentric reference distance were made to change appropriately between the presentations.

Since this discussion applies to relational cues in general, the analysis of the perceived scalar depth associated with any such cue involves three steps.

1. The rules governing the apparent position of the depth interval with respect to the egocentric reference distance should be determined. The present study is only a preliminary approach to this problem. Furthermore, it is likely that the egocentric reference distance can be determined by factors other than oculomotor cues and the specific-distance tendency. It has been demonstrated, for example, that the perceived distance of a binocular configuration and the perceived depth within that configuration are modified, despite unchanging oculomotor cues, by altering the part of the visual field in which the configuration appears (Gogel, 1964).
2. The relationship between the perceived depth interval and the *magnitude* of the egocentric reference distance should be determined for a constant value of the (relational) cue.
3. The relationship between the perceived depth interval and the *magnitude* of the (relational) *cue* should be determined for a constant egocentric reference distance.

Steps 1 and 2, but not step 3, ought to be independent of the cue being considered. Thus, considerable predictive generality could result from this type of analysis. Even without more information on step 1, it is obvious that different results would be obtained from different situations, reflecting perhaps the difficulty of obtaining consistent results in studies of the psychophysics of such relational cues to distance as binocular disparity.

In a previous article (Gogel, 1963) it was emphasized that the per-

ception of a depth interval resulting from relational cues usually could not be explained by the difference between perceptions resulting from egocentric cues. This is consistent with the assumed distinction between egocentric and exocentric cues. From the present study it can be concluded also that the scalar perception of depth from relational cues often involves a calibration process using the egocentric reference distance. It is asserted that the interrelation of egocentric scalar factors and relational nonscalar cues is basic to the production of a visual world in which relational cues often are not identical to absolute cues but in which perceived extents characteristically are scalar.

Notes

This investigation was supported by PHS Research Grant NS-08883 from the National Institute of Neurological Diseases and Strokes. The author thanks Robert E. Newton and Jerome Tietz for their help in collecting and analyzing the data. Revision received for publication August 31, 1971.

1. Equations 1 and 2 suggest that the calibration of a binocular disparity by a perceived egocentric distance is direct. Another possibility is that the perception of egocentric distance determines the perceived size (S') per unit of retinal size (θ) of frontoparallel planes in the vicinity of the disparity and that this S'/θ in turn affects the perceived depth from binocular disparity (Gogel, 1964). In either case, the perception of egocentric distance is basic to the perception of depth from a binocular disparity.

References

- Biersdorf, W. 1966. Convergence and apparent distance as correlates of size judgments at near distances. *Journal of General Psychology* 75: 249-264.
- Biersdorf, W., Ohwaki, S., and Kozel, D. J. 1963. The effects of instructions and oculomotor adjustments on apparent size. *American Journal of Psychology* 77:1-17.
- Crannell, C. W., and Peters, G. 1970. Monocular and binocular estimations of distance when knowledge of the relevant space is absent. *Journal of Psychology* 76:157-167.
- Epstein, W., and Landauer, A. A. 1969. Size and distance judgments under reduced conditions of viewing. *Perception and Psychophysics* 6:269-272.
- Fillenbaum, S., Schiffman, H. R., and Butcher, J. 1965. Perception of off-sized versions of a familiar object under conditions of rich information. *Journal of Experimental Psychology* 69:298-303.
- Foley, J. M. 1967. Binocular disparity and perceived relative distance: An examination of two hypotheses. *Vision Research* 7:655-675.
- Foley, J. M. 1968. Depth, size, and distance in stereoscopic vision. *Perception and Psychophysics* 3:265-274.
- Gogel, W. C. 1961a. Convergence as a cue to absolute distance. *Journal of Psychology* 52:287-301.

- Gogel, W. C. 1961b. Convergence as a cue to the perceived distance of objects in a binocular configuration. *Journal of Psychology* 52:303-315.
- Gogel, W. C. 1962. The effect of convergence on perceived size and distance. *Journal of Psychology* 53:475-489.
- Gogel, W. C. 1963. Visual perception of size and distance. *Vision Research* 3:101-120.
- Gogel, W. C. 1964. Perception of depth from binocular disparity. *Journal of Experimental Psychology* 67: 379-386.
- Gogel, W. C. 1968. The measurement of perceived size and distance. In *Contributions to sensory physiology*, vol. 3, ed. W. D. Neff. New York: Academic Press.
- Gogel, W. C. 1969a. The absolute and relative size cues to distance. *American Journal of Psychology* 82: 228-234.
- Gogel, W. C. 1969b. The effect of object familiarity on the perception of size and distance. *Quarterly Journal of Experimental Psychology* 21: 239-247.
- Gogel, W. C. 1969c. The sensing of retinal size. *Vision Research* 9:3-24.
- Gogel, W. C., Hartman, B. O., and Harker, G. S. 1957. The retinal size of a familiar object as a determiner of apparent distance. *Psychological Monograph* 71 (no. 442).
- Gogel, W. C., and Mertens, H. W. 1967. Perceived size and distance of familiar objects. *Perceptual and Motor Skills* 25:213-225.
- Gogel, W. C., and Mershon, D. H. 1969. Depth adjacency and simultaneous contrast. *Perception and Psychophysics* 5:13-17.
- Gogel, W. C., and Newton, R. E. 1969. Perception of off-sized objects. *Perception and Psychophysics* 5:7-9.
- Heinemann, E. G., Tulving, E., and Nachmias, J. 1959. The effect of oculomotor adjustments on apparent size. *American Journal of Psychology* 72:32-45.
- Kinchla, R. A. 1971. Visual movement perception: A comparison of absolute and relative movement discrimination. *Perception and Psychophysics* 9:165-171.
- Ono, H. 1970. Some thoughts on different perceptual tasks related to size and distance. In *Human space perception: Proceedings of the Dartmouth Conference*, ed. J. C. Baird, ch. 3. *Psychonomic Monograph Supplements* 3(13, whole no. 45).
- Richards, W., and Miller, J. F. 1969. Convergence as a cue to depth. *Perception and Psychophysics* 5:317-320.
- Schiffman, H. R. 1967. Size estimations of familiar objects under informative and reduced conditions of viewing. *American Journal of Psychology* 80:229-235.
- Schober, H. 1954. Über die Akkomodations Rugelage. *Optik* 11:282-290.
- Von Kries, J. 1925. Notes on perception of depth. In *Treatise on physiological optics*, by H. V. Helmholtz, vol. 3, trans. and ed. J. P. C. Southall, pp. 396-400. Menasha, Wisc.: Banta.

continued from page 617

- Clinebell, Howard J. *The People Dynamic: Changing Self and Society through Growth Groups*. New York: Harper & Row, 1972. Pp. x, 176. \$4.95.
- Crosby, Andrew. *Creativity and Performance in Industrial Organization*. New York: Harper & Row, 1972. Pp. xiii, 197. Paperback, \$3.75.
- Cull, John G., and Colvin, Craig R. *Contemporary Field Work Practices in Rehabilitation*. Springfield, Ill.: Charles C Thomas, 1972. Pp. xvii, 323. \$16.75.
- Cull, John G., and Hardy, Richard E., eds. *Vocational Rehabilitation: Profession and Process*. Springfield, Ill.: Charles C Thomas, 1972. Pp. ix, 547. \$18.50.
- Dahlstrom, W. Grant; Welsh, George Schlager; and Dahlstrom, Leona E. *An MMPI Handbook*. Vol. 1, *Clinical Interpretation*. Minneapolis: University of Minnesota Press, 1972. Pp. xxvi, 507. \$18.75.
- Dedek, John F. *Human Life: Some Moral Issues*. New York: Sheed & Ward, 1972. Pp. 180. \$5.95.
- Deese, James. *Psychology as Science and Art*. New York: Harcourt Brace Jovanovich, 1972. Pp. viii, 120. Paperback, \$2.95.
- DuBrin, Andrew J. *Women in Transition*. Springfield, Ill.: Charles C Thomas, 1972. Pp. ix, 179. \$11.75; Paperback, \$6.75.
- Edwards, Allen L. *Experimental Design in Psychological Research*. 4th ed. New York: Holt, Rinehart and Winston, 1972. Pp. xvii, 488.
- Eissler, Ruth, S.; Freud, Anna; Kriss, Marianne; Lustman, Seymour L.; and Solnit, Albert J., eds. *The Psychoanalytic Study of the Child*. New York: Quadrangle Books, 1971. Pp. viii, 616. \$12.50.
- Ekman, Paul; Friesen, Wallace, V.; and Ellsworth, Phoebe. *Emotion in the Human Face*. New York: Pergamon Press, 1972. Pp. xii, 187. \$8.50.
- Ellis, Henry C. *Fundamentals of Human Learning and Cognition*. Dubuque, Iowa: Wm. C. Brown, 1972. Pp. xix, 238. Paperback, \$3.95.
- Eysenck, H. J., ed. *Readings in Extraversion-Introversion*. Vol. 1, *Theoretical and Methodological Issues*. Vol. 2, *Fields of Application*. Vol. 3, *Bearings on Basic Psychological Processes*. New York: Wiley, 1972. Pp. 416, 355, 640. \$14.95, \$13.95, \$13.95; as a set, \$41.50.
- Fait, Hollis F. *Special Physical Education: Adapted, Corrective, Developmental*. Philadelphia: W. B. Saunders, 1972. Pp. xi, 442. \$9.25.
- Fingarette, Herbert. *The Meaning of Criminal Insanity*. Berkeley and Los Angeles: University of California Press, 1972. Pp. viii, 271. \$10.00.
- Foulds, G. A. *Personality and Personal Illness*. New York: Harper & Row, 1972. Pp. xi, 344. Paperback, \$3.00.
- Fraisse, Paul, ed. *L'année psychologique*. Paris: Presses Universitaires de France, 1972. Pp. 347, 665.
- Fredericksen, Hazel, and Mulligan, R. A. *The Child and His Welfare*. 3d ed. San Francisco: W. H. Freeman, 1972. Pp. x, 434. \$8.50.
- Geldard, Frank A. *The Human Senses*. 2d ed. New York: Wiley, 1972. Pp. xi, 584.
- Giglioli, Pier Paolo, ed. *Language and Social Context*. Baltimore: Penguin Books, 1972. Pp. 391. Paperback, \$3.50.

continued on page 534

THE ROLE OF ASSOCIATIONS IN VERBAL-DISCRIMINATION LEARNING

Joel Zimmerman, John J. Shaughnessy, and Benton J. Underwood

Northwestern University

The two general conclusions from five experiments on the role of intrapair associations in verbal-discrimination learning were, first, that the frequency of implicit associational responses has minimal effect on verbal-discrimination and two-category classification learning, and second, that the intrapair associations which do develop in verbal-discrimination learning are not critically involved in the making of those discriminations.

The studies reported here were initiated by two developments. First is the finding that if pairs of strong natural associates (the pair king-queen, for example) are used in a verbal-discrimination list, performance does not differ appreciably from performance on a list of nonassociated pairs. Different investigators in different laboratories have not always produced results in complete agreement with the above generalization, but it is the consensus (Eberlein and Raskin, 1968; Fulkerson and Kausler, 1969; Lovelace and Schulz, 1971; Palermo and Ullrich, 1968).

That finding does not support the frequency theory of verbal-discrimination learning (Ekstrand, Wallace, and Underwood, 1966) that has in its combined form included implicit associational responses as one source of frequency. The theory originally assumed that when a common word is used in a learning task, it is likely to produce implicitly a verbal associate, an implicit associational response (IAR) with a situational frequency that is in some sense the same as the frequency produced by the word actually presented. If, then, highly associated words ('high associates') are paired in a list, one being right (R) and the other wrong (W), and if each produces the other implicitly, it should be difficult to establish a frequency differential between the two and learning should be impaired. As noted above, this expectation has not been supported. Why does the combined IAR/frequency theory fail at this point?

There is another relevant fact in this first category. If high associates are used in a list but are used as members of *different* pairs, with one

being right and the other wrong (crossed associates), learning is impaired somewhat (Eberlein and Raskin, 1968; Fulkerson and Kausler, 1969). This may suggest that implicit associational responses are occurring as assumed by the theory. But if so, why is there not also the expected impairment when the associates are members of the same pair? One possibility is that an association between words in a pair is in some way involved in a positive manner with the establishment of the discrimination. This leads to the second development that has prompted the present studies.

It is a well-replicated fact that as a verbal-discrimination list of unrelated pairs is learned, an association develops between the pair members (see, for example, Spear, Ekstrand, and Underwood, 1964). Does this association have any functional—hence, theoretical—significance for the acquisition of the verbal discrimination per se? Some investigators are inclined to believe that it does, that in some way the association aids in identifying which member of the pair is right and which is wrong (Fulkerson and Kausler, 1969). The lack of an explicit articulation of the mechanisms involved makes it difficult to test the notion, but the seeming inevitableness of the development of associations between the pair members makes it difficult to dismiss the idea that those associations may play a fundamental role in establishing the discriminations.

The experiments to be reported may be viewed as attempts to identify the function (if any) that associations play in verbal-discrimination learning—as attempts toward a better understanding of the failure of the combined IAR/frequency theory. The first two experiments deal with natural associates, the subsequent ones with associations established in the laboratory. Experiment I merely establishes the fact that the lack of an effect of paired natural associates on verbal-discrimination learning could be replicated in our laboratory.

EXPERIMENT I

Method

—Materials—Four different lists were used, each consisting of 16 pairs. Two of the lists were made up of pairs of *highly associated words* as determined by the associative frequencies in the Palermo-Jenkins (1964) norms for college students. The associates varied from 41% to 71% in frequency in the forward direction. The backward associations were in some cases of relatively low frequency, but it is doubtful that this is of moment when the words appear together in a list. Two other lists had pairs of *words with zero frequency* in the norms. For each type of list, the associates, high or zero,

were *paired* for one list and *crossed* for the other. It probably would not have been necessary to use a crossed list for the zero associates.

Since the words were also used in Experiment II, they are presented in Table 1. Of course, for the lists of high associates, one member of the pair was a W word and the other the R word, these functions remaining the same in the paired and crossed lists.

Table 1. Words used in Experiments I and II

High associates		Zero associates	
bitter-sweet	hammer-nail	bath-cry	justice-bird
black-white	king-queen	cold-food	lion-white
bread-butter	long-short	deep-trucks	mountain-short
chair-table	needle-thread	finger-mouse	rough-star
slow-fast	blossom-flower	guns-queen	bed-sweet
fruit-apple	eagle-bird	hard-sleep	carpet-butter
girl-boy	bed-sleep	head-grass	city-flowers
green-grass	lamp-light	joy-mind	doctor-table

—Procedure—Each of the four lists was presented to independent groups of 20 subjects each, for one study trial and five anticipation trials. The rate was 2:2 sec, with a 4-sec intertrial interval. The subjects, who were instructed not to guess on the study trial, were 80 college students assigned to groups by a schedule that block-randomized the four lists.

Results

Since there was no apparent interaction in the error scores across trials for the four lists, the basic measure used was mean total number of errors for the five anticipation trials. The two zero lists had means of 8.40 and 8.20 for the paired and crossed lists respectively, and the standard deviations were 8.24 and 6.00. The means for the lists of high associates were 9.05 when the associates were paired and 13.75 when crossed, with the standard deviations being 7.30 and 6.78.

Analysis showed that no source of variance approached significance. The largest *F*, for zero associates versus high associates, was only 1.97. The data thus confirm the general finding that paired high and paired low associates in verbal-discrimination lists do not differentially influence the learning. The results were in the appropriate direction to support the other general finding (that crossed associates do impair the learning), but the difference lacked statistical significance.

EXPERIMENT II

The data in Experiment I would seem to indicate that the combined IAR/frequency theory was again not supported. High associates,

whether paired or crossed in a verbal-discrimination list, did not retard learning. One issue left in doubt is whether or not implicit associational responses were occurring as assumed by the theory. This issue is related to the general problem of the control a subject has over his encoding of responses, of which the implicit associational response is one type. Or to put it bluntly, if a subject can detect that his implicit associational responses will produce difficulties in learning, can he 'turn them off'? Cramer (1969) has concluded that a subject does have such control, but a conception such as that held by Wickens (1970) would deny this possibility. A direct attack on the problem seems difficult to devise, but the conditions of Experiment II were related to this matter. We set about to see if we could find a discrimination task in which differences in learning between high and zero associates would emerge.

There were four conditions in the experiment. Palermo and Ullrich (1968) reported a study in which words were presented individually, with the subject required to make different responses to the associated words. For example, when the word king was shown, the subject had to remember to pronounce it, but not to pronounce the word queen when it was shown. Thus, it was a *two-category classification task*. Their results showed that a list of high associates was more difficult to learn than a list of low associates. Unfortunately, a mixed list (both high and low associates were in the list) was used, which could influence performance for reasons other than differential learning difficulty.

In any event, the first condition in Experiment II utilized the Palermo-Ullrich approach. The second condition was like the first, but each time a word was presented its associate was *also* presented, in parentheses. This can be viewed as an attempt to produce the implicit associational response for the subject. The third condition was a mixed-task list in which half the items were constituted of verbal-discrimination pairs and half were presented singly as in the first two conditions. A fourth condition involved verbal-discrimination lists with high and zero associates. This was included to make sure that with new subjects we could replicate the findings of Experiment I.

The intent behind the use of the classification task should not be misunderstood. No assumption was made that the processes of learning or the strategies of learning which underlie the classification task are the same as those which underlie verbal-discrimination learning. Rather, the results reviewed above suggest that implicit associational responses may not occur (or may occur with very low frequency) in the verbal-discrimination task, where pairs are used. So the classifica-

tion task in the first condition presented items individually, to enhance the spontaneous occurrence of their associates, and in the second condition actually presented the associates (in parentheses), as a means of assuring their presence, if not their spontaneous occurrence. If differences in learning lists of high and low associates did not emerge under these conditions, it could be taken as evidence that the role of implicit associational responses has been overestimated.

Method

—Materials and conditions—Each of the four conditions involved two lists, one of high associates and one of zero associates. The words used were the first 12 pairs of both groups of 16 in Table 1. In the first condition (*condition S*), the 24 words were presented singly, with the subject required to learn to say yes to 12 of the words and no to the other 12. This was true for both the high associates and for the zero associates. For the list of high associates, one member of each pair was randomly determined to be a 'yes' word, the other member a 'no' word. For the zero associates, the yes and no assignment was determined randomly for the two members of each pair as given in Table 1. The second, *S(A) condition* was exactly the same as the first, except that the associated word (or the zero-associate pair member for the zero list) always occurred in parentheses below the word for which the discrimination was required. For example, when the word king was shown, the word queen occurred below it in parentheses, and when 'queen' appeared, 'king' occurred below it in parentheses. The same was true for the arbitrary pairs in the zero list.

The third, *S-VD condition* was the transition condition (between classification learning and verbal-discrimination learning), in which the list consisted of six verbal-discrimination pairs plus 12 words presented singly as in condition S. The 12 pairs were arbitrarily divided in half and the first six pairs occurred as verbal-discrimination pairs, the second six pairs as 12 individually presented words. For half the subjects in this condition, the item functions were reversed. The fourth, *VD condition* involved the learning of the two 12-pair verbal-discrimination lists, one of high associates and one of zero associates.

—Procedure and subjects—For conditions S and S(A) the 24-item lists were presented as though they were paired-associate lists with only two response terms, yes and no. A 2:2-sec rate was used in that the word was presented for 2 sec during which the subject was to respond yes or no, whereafter there followed a 2-sec period during which the shutter lifted on the memory drum to show the correct response (yes or no). Four different random orders were used, the only restriction being that associated words not occupy adjacent positions. The intertrial interval was 4 sec.

The instructions stressed that the assignment of a yes or a no was quite arbitrary and did not imply any system. In condition S(A) the subject was further told that the word in parentheses might help him learn, since its correct response was always 'opposite' of the current correct response. For

condition S-VD the subject was instructed as in condition S for the single words and for the paired words was given the usual verbal-discrimination instructions. The rate remained at 2.2 sec, with the correct member of each verbal-discrimination pair being shown when the shutter lifted. It may be noted that for conditions S and S(A) the total time per trial (excluding the 4-sec intertrial interval) was 96 sec; for condition S-VD the time per trial was 72 sec. Condition VD was straightforward verbal-discrimination learning of the 12 pairs and hence involved 48 sec per trial.

In summary, there were four conditions, differentiated by the method of presenting the material. For each condition there was always a list consisting of the 12 high associates and always a second list consisting of the 12 pairs of zero associates. Each list within each condition was given to a different group of subjects; hence eight groups were required. Each group consisted of 24 college students, assigned to a particular list and condition by a block-randomized schedule. For condition S-VD, the randomization also included the particular list form, each form for both the high and zero lists being assigned to 12 subjects.

All lists were presented for an initial study trial followed by 12 anticipation trials, with the subject instructed to respond to all items after the study trial. After that, all subjects were, unexpectedly, asked to recall all of the words they could, writing them in any order they wished on the prepared data sheet.

Results

There is little background literature on the task used in *conditions S and S(A)*. It may be viewed as a 24-pair, paired-associate list with only two response terms. Our preference is to view it as a two-category classification task without any systematic basis for the classification. The acquisition curves are plotted in Figure 1 in terms of mean errors per trial. It is quite evident for both conditions that the (singly presented) high associates did not result in performance inferior to that for the zero associates; in fact, the differences favored the high associates on the initial trials.

The second fact to be noted is that learning was more rapid under condition S than S(A); the presence of the words in parentheses in condition S(A) retarded learning. Because of the suggestive interactions with trials seen in Figure 1, the statistical analysis included trials as one factor. Only one source of variance (in addition to trials) was significant, namely, the difference between conditions S and S(A) [$F(1, 92 = 5.46, p < .05)$].

If the associated words in these two conditions were somehow being learned via the natural association between members of a pair, then the numbers of errors made on each of the associated words in a pair should have been more comparable in frequency than the numbers for

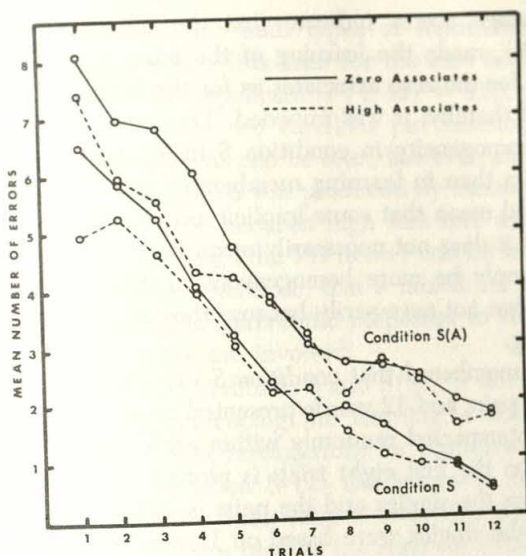


Fig. 1. Mean number of errors by trials for two-category classification tasks as a function of associations between pairs; conditions S and S(A), Experiment II

words in pairs with no natural associates, as in the zero pairs. For each subject a determination was made of the difference in the numbers of errors he made on the two words of each associated pair. For example, the numbers of errors he made on 'king' and on 'queen' across 12 trials were noted, and the difference between these frequencies was recorded. For each subject there were 12 such differences. As a control, the same differences were determined for the 'artificial' pairs of the zero lists. Since number of total errors differed, each subject's score was expressed as a percentage of total errors made. A low percentage score would mean that a subject had made highly comparable (more homogeneous) numbers of errors on the two words in a pair.

For condition S the means of these percentage scores were 69.3% and 54.7% for the zero pairs and the associated pairs respectively; the corresponding values were 36.1% and 41.7% for condition S(A). Following an arcsin transformation, an analysis of variance was performed. All F s were based on 1 and 92 df . The zero versus high comparison was not significant ($F = 1.55$). However, the difference between conditions S and S(A) was highly significant ($F = 25.24$), and the interaction ($F = 6.14$) was just short of the .01 significance level.

The percentage scores indicate that the parenthetical words, in condition S(A), made the learning of the pairs highly homogeneous—as much so for the zero associates as for the high. Yet, learning was not facilitated thereby; it was impeded. They also indicate that there was greater homogeneity in condition S in learning the members of associated pairs than in learning members of pairs with zero association. This could mean that some implicit associational responses were occurring, but it does not necessarily mean that. Words that are associated may simply be more homogeneous in difficulty than nonassociated words, but not necessarily because they are eliciting each other during learning.

It will be remembered that *condition S-VD* consisted of six verbal-discrimination pairs and 12 words presented singly. The pairs and the singles were interspersed randomly within each trial. The mean number of errors on the first eight trials is plotted in Figure 2. The difference between the singles and the pairs is of no consequence, since the scores for the singles were based on 12 possible errors, and those for the pairs on six possible errors.

Figure 2 suggests that with this mixed list, the level of association

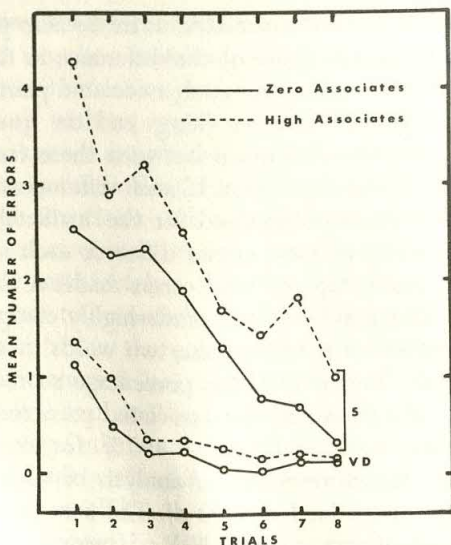


Fig. 2. Mean number of errors by trials for combined two-category classification (S) and verbal-discrimination (VD) tasks as a function of associations between pairs; condition S-VD, Experiment II

was producing an effect; for both types of items, the errors were greater for the high-associate lists than for the zero lists. Statistically, the zero/high variable was significant [$F(1, 46) = 4.53, p < .05$], with the interaction between the two variables just reaching significance [$F(1, 46) = 4.18, p < .05$]. As can be seen, however, almost the entire effect of the level of association was produced by the single (S) items. Tested alone, the difference between high and zero associates in the verbal-discrimination learning (the VD items) was far from significant. The outcome seems to be, therefore, that a mixed list produces a retardation in the learning of differential responses to singly presented items when high associates are involved.

As was pointed out for conditions S and S(A), if the high associates presented singly were influencing the learning, a concomitant phenomenon should have been homogeneity in error frequency for the associated words. In condition S-VD differences in error frequency for each associated pair for each subject were also determined, with the sum of the six differences used as the score. A similar distribution of 24 scores was determined for the nominal pairs of the zero associates presented singly. For subjects having the high associates, the mean was 51.0%; for those with zero associates, 73.4%. Following an arcsin transformation, the difference was found to be highly significant [$F(1, 46) = 10.00, p < .01$]. We may thus conclude that there was more homogeneous learning of associated than nonassociated words under condition S-VD.

Condition VD consisted of the learning of two verbal-discrimination lists, one of paired high associates and one of zero associates. It was a replication of the comparable conditions of Experiment I, using 12 pairs instead of 16. The results were the same, with learning extremely rapid for both lists. The mean number of errors across 12 trials was 6.33 for zero associates, 4.46 for high. The corresponding standard deviations were 5.83 and 5.00, and the F was 1.37. Thus again, high associates paired in a verbal-discrimination list had little effect on learning.

After the 12 learning trials under each condition, the subjects had been asked to *free recall* as many of the words as possible. An analysis of variance of the eight groups for the number recalled showed more of the high associates being recalled than the low ($F = 26.50$), the means being 20.44 and 18.01. There were no condition effects nor any interaction. Clustering of high associates was apparent in all conditions. For condition S the zero associates did *not* cluster by 'yes' and 'no' categories, and no other basis for clustering was apparent. In condition

S(A) two-word clustering of the zero associates was heavy, being based on the two words that always occurred together (one in parentheses). Of the total words recalled, 76% were recalled as pairs.

The basic conclusion from Experiments I and II is that natural associates have very little influence on discrimination learning. This was true for verbal-discrimination lists and for two-category classification learning of words presented singly. It could be argued, of course, that high associates in the classification task produced both positive and negative effects essentially equal in magnitude. For example, when 'king' was presented and the correct response shown, the subject could have practiced the response both to 'king' and to 'queen' ('king,' yes; 'queen,' no). But it would seem that the parenthetical words would facilitate such rehearsal, and yet learning was slower in condition S(A) than in condition S. If there are balancing effects that lead to no differences in learning lists of high and low associates, they were not apparent. High associates had no effect on classification learning and no effect on verbal-discrimination learning. The evidence, therefore, does not support theorizing that makes the association between words in verbal-discrimination lists a fundamental mechanism in the learning. The only substantial negative effect of associates was found in classification learning when this was interspersed with verbal-discrimination learning. This finding will be evaluated in the general discussion.

EXPERIMENT III

The plan of Experiment III was to develop associations between unrelated words by *paired-associate learning* and then to use these pairs in verbal-discrimination lists. If the association is a critical part of the discrimination required in verbal-discrimination learning, performance on that discrimination task should be positively related to (improved by) the number of trials on the paired-associate task.

The design of Experiment III was directed to a certain extent by the assumption that the usual verbal-discrimination learning is based primarily on a frequency discrimination between the members of each pair. Thus the learning of paired associates would produce not only associations but also a high level of situational frequency for the individual words. Insofar as Weber's law holds for frequency discriminations, this base frequency could then *retard* verbal-discrimination learning, here also in direct relation to the number of trials on the paired-associate task.

Our concern with this latter point might have been minimized had we known of the recently published work of Hintzman and Block (1971), which shows that apparent frequency can be highly specific to a given point in time even for the same type of task. However that may be, additional conditions were devised which would allow frequency inputs corresponding to those in paired-associate learning but would not allow associations to develop. There were five conditions, three involving different degrees of paired-associate learning (0, 8, 16 trials) and two corresponding to 8 and 16 paired-associate trials but without associative learning. All subjects learned the same verbal-discrimination list as the final task.

Method

—Materials and conditions—Thirty two-syllable concrete nouns were used to form 15 pairs for the paired-associate and verbal-discrimination lists. The words varied in frequency from 20 to 40 in the Thorndike-Lorge count. The paired-associate learning was bidirectional in that across trials each word in a pair was used equally often as a stimulus term and as a response term. The subject was instructed that whenever a word appeared he was to try to give the other word in the pair. In one group (*condition PA-8*) the list was presented for 8 anticipation trials (including the study trial); in another (*condition PA-16*), for 16 trials. Four different orders of the pairs were used, with an intertrial interval of 4 sec.

Two other groups (*conditions F-12 and F-24*) were exposed to the words in such a way as to approximately match the input frequencies for groups PA-8 and PA-16 but at the same time to minimize the possibilities of associative formation between or among words. In paired-associate learning, it was assumed, each word would gain three frequency inputs every two trials, two from exposure frequency and one from a correct anticipation or rehearsal. Therefore, the induced frequencies for the F groups were 12 and 24 to match the input frequencies for groups PA-8 and PA-16 respectively. (It should be noted that there was no need to have frequencies exactly equivalent in the two tasks before verbal-discrimination learning, since the base frequencies in both cases would be so high that frequency differences within each task would be nondiscriminable.) The pairs for groups F-12 and F-24 were presented exactly the same as in the anticipation procedure for groups PA-8 and PA-16, but these F-group pairs were never the same from trial to trial and the F-group subject was explicitly instructed not to form associations but rather to learn the *frequencies* with which each word was presented. Twelve different random pairings were used (including a word paired with itself), subject to the restriction that in every 12 trials each word occur a minimum of 10 times and a maximum of 14 times. Group F-12 was given 12 exposure trials; group F-24 got 24 trials, with each of the 12 random pairings used twice.

The subjects in groups F-12 and F-24 were told that they would be tested for frequency knowledge periodically. A test was given after every third

trial, by presenting six of the words on a sheet of paper, with the request that these words be rank-ordered for the frequency with which they had been presented. A different set of six words was used on each test, so that for group F-12, 24 of the 30 words were tested. The same tests were used for the second block of 12 trials for group F-24. These tests were used merely to make the subjects attend to the words, and no results for them will be reported. The subjects in a fifth group, the *control condition*, got only the verbal-discrimination list.

—Procedure and subjects—All lists were presented at a 2:2-sec rate. Immediately after paired-associate learning, or after the frequency inputs for groups F-12 and F-24, instructions for verbal-discrimination learning were given. No mention was made of the relationships between the words in the two tasks. It should be clear that the verbal-discrimination pairs were exactly the same as the pairs in paired-associate learning. The verbal-discrimination learning continued until all responses were correct on a single trial; the first trial was a study trial (no guessing), and all subsequent trials were anticipation trials with the subject responding to every pair. Of course, the position of the two words within a pair varied from trial to trial, and there were four different orders of the pairs.

The subjects were college students, 22 in each of the five groups. They were assigned to conditions according to a block-randomized schedule.

Results

On the *paired-associate learning*, group PA-8 produced a mean of 10.73 correct responses on the eighth trial. During the eight trials, 3 of the 22 subjects had all pairs correct on a single trial at least once. Group PA-16 had 11.86 mean correct responses on the eighth trial, 13.86 on the sixteenth trial. Of the 22 subjects, 15 achieved the criterion of one perfect on at least one trial.

The criterion for *verbal-discrimination learning* was one perfect trial. Table 2 shows the mean number of trials required to reach the criterion and also the mean number of errors on the first trial. The first-trial measure does not distinguish among the groups ($F < 1$), but the trials-to-criterion measure does [$F(4, 105) = 2.83$, $p < .05$]. As can be seen in Table 2, in terms of the latter measure the two PA

Table 2. Mean number of errors on first trial and mean number of trials to criterion, Experiment III

Condition	First trial ^a	Criterion
Control	3.14	3.91
PA-8	3.41	3.00
PA-16	3.41	3.18
F-12	4.45	4.32
F-24	3.91	4.73

^aThe σ_M based on within-groups variance was .47 for the first trial and .44 for the criterion measure.

groups were superior to the control; the two F groups, inferior. However, the control group did not differ from the two PA groups combined [$F(1, 105) = 2.34, p > .05$], which indicates that the associations developed in paired-associate learning had no appreciable effect on verbal-discrimination learning.

The two F groups combined were inferior to the two PA groups [$F(1, 105) = 10.76, p < .01$]. Two factors may be involved in this difference. First, if there was transfer of 'learning to learn' from the paired-associate to the verbal-discrimination task, the PA groups would have been facilitated relative to the F groups (and the small superiority of the PA groups over the control may also have been due to that transfer of 'learning to learn'). Secondly, the F groups were actually making frequency judgments in their first task, and the accumulated frequencies may have carried over somewhat to their verbal-discrimination task, thereby making those later frequency discriminations more difficult. (We presume the differences between their first and second tasks prevented such carryover for the PA groups.) It might be argued that a third fact should be considered, namely, that associations were developed by the first task for the PA groups and not by that for the F groups. The complete lack of a relationship between degree of paired-associate learning and verbal-discrimination learning would deny this, but the matter needs further consideration.

The data in Table 2 indicate that degree of paired-associate learning was irrelevant to verbal-discrimination learning. Two additional tests were made of this potential relationship. For each subject in groups PA-8 and PA-16, we determined the seven items that had been given correctly the greatest number of times in paired-associate learning, and the seven anticipated correctly the fewest number of times. This division represents a within-subjects difference in degree of paired-associate learning. We then asked about the errors made on the corresponding pairs of the two classes in verbal-discrimination learning. There was no relationship ($F_s < 1$).

As a second test, the correlations between the two learning tasks were determined. If degree of learning of paired-associates is important for verbal-discrimination learning, those subjects doing best on paired-associates learning should have done best on verbal-discrimination learning. Of course, insofar as there is some general learning ability that might underlie performance on both tasks, we would expect the relationship to be positive. However, the correlations were .38 for PA-8, and .32 for PA-16. Neither attained the 5% significance level of .42.

It seems, therefore, that the conclusion from Experiment III must be

the same as that from Experiments I and II. Associates developed in the laboratory 'behave' very much like natural associates; neither kind is of much consequence for verbal-discrimination learning. The final two experiments will give further support to this conclusion.

EXPERIMENTS IV AND V

As was pointed out in the introduction, many studies of verbal-discrimination learning have shown that associations develop between the pair members as performance increases over trials. The present studies have asked whether this association is critically involved in the development of the discrimination or whether the association is irrelevant. Experiments I-III produced data which have been interpreted to mean that the association is irrelevant. In Experiments IV and V, the approach was to *prevent* associations from forming, on the grounds that if the association is critically involved in verbal-discrimination learning, preventing the formation of the associations should at least retard that learning.

In Experiment IV three conditions were employed. One was standard verbal-discrimination learning. The second condition involved re-pairing of R and W words on each trial; and in the third condition, new W items were used on each trial. Frequency theory per se assumes that the association is of no consequence for the discriminations and makes specific predictions about the relative rate of learning to be expected for the three conditions.

According to the theory, the re-pairing of the R and W items in the second condition should be of no consequence, since the frequency differential between R and W words should, on the average, be as discriminable with constantly changing pairs as with constant pairs. Dominowski (1966) has reported that re-pairing of R and W items had no influence on verbal-discrimination learning; hence the first two conditions may be viewed as a replication of his work. It is also known that re-pairing of R and W items after considerable verbal-discrimination learning has occurred will produce a considerable disturbance in performance (Kanak and Dean, 1969). This problem, to which Wallace and Nappe (1971) have addressed themselves, is not necessarily related to the issues involved in the present procedure, wherein the re-pairing occurred on every trial from the beginning of learning.

The third condition, in which new W words are inserted on each trial, should produce very rapid discrimination. It is essentially a two-choice recognition situation in which the W words are always new,

hence have zero situational frequency, while the frequency of the R word is increasing steadily over trials. Yelen (1969) found very great facilitation when he added new W items on each trial, even in a four-choice situation with nonsense syllables. The effect should be even greater when the more common two-choice procedure is employed.

Experiment V repeated the first two conditions of Experiment IV, with a change in the instructions to the subjects. This was done in an effort to clarify one of the findings of Experiment IV.

Method

—Materials and conditions—The lists consisted of 15 pairs. Initially, 90 words were chosen from the list provided by Paivio, Yuille, and Madigan (1968). All words had ratings of concreteness of 6.0 or higher, and all contained ten letters or less. From the pool, 15 words were randomly chosen to serve as the R words for all conditions. Five other randomly formed groups of 15 words served as the W words for the condition in which new W words were inserted on every trial (*condition N*). For the usual or constant condition (*condition C*) five lists were used. In all of these the R words were the same but the W words differed across subjects to correspond to the five 15-word sets used for condition N. Each list was given to an equal number of subjects. These same five lists were used for the condition in which the R and W words were re-paired on each trial (*condition R*). In Experiment V, only conditions C and R were run.

—Procedure and subjects—The lists were presented for one study and four anticipation trials at a 2:1-sec rate, with a 3-sec intertrial interval. A different order of the pairs was used on each trial, and the position of the two words (above or below the other) during the anticipation interval changed unsystematically over trials. The correct word was shown alone during the feedback interval. For Experiment IV the usual verbal-discrimination instructions were given. For Experiment V the instructions were amplified for condition R, where the subjects were told that the words would be paired differently on each trial but that a given correct word would always remain the correct word.

A total of 90 subjects was used in Experiment IV, 30 being assigned to each condition. In conditions C and R, 6 subjects were assigned to each of the five list forms. The three conditions were randomized in blocks of 15, which included 1 subject for each of the five forms for conditions R and C, and 5 subjects for condition N. The subjects were assigned to conditions in order of appearance.

In Experiment V, in which only conditions R and C were used, 20 subjects were assigned to each condition. Only one list was used. Other than the number of subjects in the two conditions, and the use of a single list, the only difference between Experiments IV and V was in the instructions to the subjects.

Results

The results of condition N of Experiment IV may be quickly disposed of. In this condition new W items were used on each trial. On the first anticipation trial, the mean number of errors was 1.00. Of the 30 subjects, 17 had perfect performance on the first anticipation trial, 2 made one error each on the second trial, 1 made one error on the third, and all had perfect performance on the fourth trial

The results for conditions C and R for both experiments are shown in Figure 3. For Experiment IV (left panel) the performance under condition R was inferior to that under condition C. In terms of mean number of errors across all trials, the difference was statistically significant [$F(1, 58) = 4.42, p < .05$]. The largest difference appeared on the first trial. It seems unlikely that a re-pairing decrement would be so large on the first trial, when the associations would be relatively weak, and rather more likely that the difference was due to the subject's being puzzled by the unexpected appearance of new pairings. Of course, for the subject to be puzzled would require that some associations had been developed between the pairs on the study trial—otherwise, the change would go unnoticed.

However this may be, the purpose for conducting Experiment V was to eliminate, by instructions, this possibility of confusion for condition R. As a consequence, it seems, the difference between the two conditions was eliminated, and Dominowski's findings (1966) were confirmed.

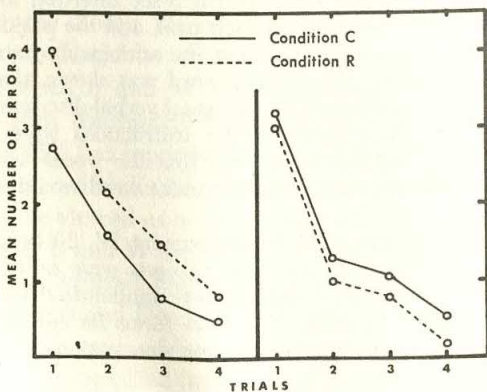


Fig. 3. The effect of the re-pairing of right and wrong members of pairs in verbal-discrimination learning; condition C is the control, condition R the re-pairing condition; Experiment IV (left panel), Experiment V (right panel)

This is seen in the right panel of Figure 3. Actually, performance was a little better under condition R than C, but the difference was far from significant. A comparison of the two panels shows that the major change in performance between the two experiments was in condition R, where the reduction in mean total number of errors was 3.50 [$t(48) = 2.22$, $p < .05$]. The difference for condition C was .57 errors ($t = .35$).

In condition R of Experiment V the instructions about the changes in the pairings should have eliminated any tendency for a subject to try a strategy involving associations between the members of the pairs. In this sense, then, the data indicate again that such associations are not a necessary part of verbal-discrimination learning. Both Experiments IV and V indicate that the prevention of development of associations between members of pairs does not retard learning; and both, in a positive sense, give strong support to the frequency theory (as opposed to the combined IAR/frequency theory) of verbal-discrimination learning.

GENERAL DISCUSSION

The bulk of the evidence in the five experiments leads to the conclusion that associations between members of pairs are not critically involved in verbal-discrimination learning. In Experiments I and II the pairing of strong natural associates had no influence on learning. In Experiment III the development of associations between neutral words did not influence subsequent performance. And in Experiment V prevention of the development of associations had no effect if the subject was appropriately instructed.

Yet, having reached the conclusion that associations are of little consequence, the question remains as to why strongly associated words do not retard verbal-discrimination learning. If two words are paired, if they are strongly associated, and if they persistently elicit each other, how can a frequency differential develop—or indeed, any other stable differentiating cue?

We believe the evidence points to the likelihood that the assumption of *persistent* elicitation is incorrect. If that assumption were true, then the classification task used in condition S of Experiment II would almost certainly have been influenced by the associates. It does not seem possible that the null effects obtained would have been found if the associates were in fact persistently eliciting each other implicitly. In other words, it may not be a matter of the subject's not recognizing that

highly associated words are present; the heavy clustering in free recall would probably indicate that he is recognizing them. It seems more a matter of the subject's being capable of ignoring the associations, of turning them off. At least two problems confront this position.

First, in condition S-VD of Experiment II, where classification learning and verbal-discrimination learning were inserted in the same task, the high associates did indeed retard learning, primarily the two-category classification learning. We have no adequate accounting of this other than to say that each task may well require different learning strategies (the use of different attributes) to develop the discriminations and that combining the tasks might have created conditions under which the subjects could not handle—prevent the occurrence of—some implicit associational responses in the classification learning.

The second problem concerns the finding by some investigators (noted in the introduction) that crossed associates produce some difficulty in learning. In the present Experiment I, the effect was present but not significant statistically. The assumption that this effect is produced by implicit associational responses may be in error. What may be more relevant here is Wallace's idea (1972) that the growth of apparent frequency as a function of input frequency is far more uniform for highly associated words than for randomly paired words.

It is known that words differ rather widely in apparent frequency as a function of constant input frequency; what Wallace assumes is that associated words are more like each other in the relationship between input and apparent frequency than are nonassociated words. This does not mean that for all pairs of associated words the growth in apparent frequency as a function of input frequency is equivalent. For example, two inputs for a given pair of associated words might be reflected in apparent frequencies of two for each word, whereas for another pair of associates the apparent frequencies might be three for each word; the relationship is the same for each word within a pair, but different between pairs.

If this assumption is correct, it means that when associates are paired in a verbal-discrimination task, the growth of frequency differences between R and W words will be consistent and stable. When these associates are crossed, the growth of apparent frequency for the W word for each frequency input may be more rapid than for the R word, making a frequency differential difficult to establish. Thus, implicit associational responses are not necessary to account for the decrement found in verbal-discrimination learning with crossed associates.

The direction of this discussion has obviously been determined by

an adherence to frequency theory. What issues would arise if it were assumed that verbal-discrimination learning is based primarily on some discriminative characteristic or cue other than frequency? For example, suppose it is assumed that verbal-discrimination learning is essentially a two-category classification problem in which the R words are stored and 'tagged' as R words, and the W words stored and tagged appropriately. If this were true, it would follow that a 12-pair verbal-discrimination list should be learned at about the same rate as a 24-item classification task with items presented singly. The data of Experiment II indicate that this is not even approximately true: in condition S (two-category classification of 24 high associates), the mean total number of errors across trials was 31.08, whereas in condition VD (the same items on a 12-pair verbal-discrimination list), the corresponding value was 4.46, even though the time per trial for the classification task was 96 sec and that for the verbal-discrimination task only 48 sec. It does not seem, therefore, that this kind of theoretical approach can meet the facts. Of course, there may be other approaches that can.

Notes

This work was supported by Contract N00014-67-A-0365-0010, Project 154-321, between Northwestern University and the Office of Naval Research. Reproduction in whole or in part is permitted for any purpose of the United States Government. Received for publication September 9, 1971.

References

- Cramer, P. 1969. Functioning of implicit associative responses in mediated transfer. *Journal of Experimental Psychology* 82:301-309.
- Dominowski, R. L. 1966. First-trial guessing and verbal-discrimination learning. *Psychonomic Science* 5:231-232.
- Eberlein, E., and Raskin, D. C. 1968. Intrapair and interpair associations in verbal discrimination learning. *Psychonomic Science* 11:145-146.
- Ekstrand, B. R., Wallace, W. P., and Underwood, B. J. 1966. A frequency theory of verbal-discrimination learning. *Psychological Review* 73: 566-578.
- Fulkerson, F. E., and Kausler, D. H. 1969. Effects of intrapair and interpair bidirectional associates on verbal-discrimination learning. *Journal of Verbal Learning and Verbal Behavior* 8:307-310.
- Hintzman, D. L., and Block, R. A. 1971. Repetition and memory: Evidence for a multiple-trace hypothesis. *Journal of Experimental Psychology* 88: 297-306.
- Kanak, N. J., and Dean, M. F. 1969. Transfer mechanisms in verbal discrimination. *Journal of Experimental Psychology* 79:300-307.

- Lovelace, E. A. and Schulz, L. S. 1971. Intrapair associations in verbal discrimination learning. *Psychonomic Science* 24:157-158.
- Paivio, A., Yuille, J. D., and Madigan, S. A. 1968. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement* 76(1, part 2).
- Palermo, D. S., and Jenkins, J. J. 1964. *Word association norms*. Minneapolis: University of Minnesota Press.
- Palermo, D. S., and Ullrich, J. R. 1968. Verbal discrimination learning as a function of associative strength between the word-pair members. *Journal of Verbal Learning and Verbal Behavior* 7:945-952.
- Spear, N. E., Ekstrand, B. R., and Underwood, B. J. 1964. Association by contiguity. *Journal of Experimental Psychology* 67:111-117.
- Wallace, W. P. 1972. Verbal discrimination. In *Human memory: Festschrift for Benton J. Underwood*, ed. C. P. Duncan and L. B. Sechrest. New York: Appleton-Century-Crofts.
- Wallace, W. P., and Nappe, G. W. 1971. Re-pairing "rights" and "wrongs" in verbal discrimination learning. *Journal of Experimental Psychology* 87:355-360.
- Wickens, D. D. 1970. Encoding category of words: An empirical approach to meaning. *Psychological Review* 77:1-15.
- Yelen, D. 1969. Verbal discrimination as a function of similarity and frequency of incorrect items. *Journal of Verbal Learning and Verbal Behavior* 8:552-558.

SOME EFFECTS OF LEARNING TO A SET OF COMPONENTS ON STIMULUS SELECTION

Jack Richardson and Sara K. Stanton

State University of New York at Binghamton

After free learning of, or learning digit responses to, a set of components (letters from a single position in consonant trigrams), the subjects learned a paired-associate list with the trigrams as stimuli. Free learning of the components had no effect on stimulus selection from the trigrams. There was no evidence that subjects changed functional stimuli to avoid negative transfer from learning different responses to the components and their corresponding trigrams. Learning the same responses to them produced positive transfer from the components to the trigrams, but no evidence of learning the responses to the additional components.

Underwood, Ham, and Ekstrand (1962) suggested that words may be selected as functional stimuli in preference to colors because subjects are more accustomed to dealing with words as stimuli. However, there has been little investigation of the effect of learning to a set of components on the subsequent stimulus selection from compound stimuli. Houston (1967) found that prior use of a component as a stimulus in a paired-associate list tended to increase the selection of that component as the functional stimulus when it was used as a component in a compound stimulus. Richardson and Chisholm (1969) showed that learning which required the use of letters from a single position in trigrams as functional stimuli increased the selection of letters from that position in subsequent learning.

Verbal material is very familiar to the usual adult subject and the units are probably well differentiated, so that learning to discriminate is not usually involved in paired-associate learning. This seems to leave three major types of learning with components that may be relevant to the subsequent selection of the components as functional stimuli from compounds. First is *familiarization*, which merely gives the subject some immediately prior experience with a set of components and serves to distinguish them from other components of the compounds. In this case, no specific response is learned to the component.

In the second type of learning, a response is learned to the com-

ponent. This should produce *positive transfer* to the learning with the compounds as stimuli if the response learned to a component is also the response to the compound containing that component. This is the positive transfer condition used in the present study. Since under this condition transfer from the component to the compound containing the component should be essentially perfect, the effect on selection from the compound is relevant to the question of what occurs to stimulus selection during overlearning. If, as James and Greeno (1967) suggest, the stimulus-selection mechanism is relaxed after paired-associate performance is perfect, then prior learning of the response to a component should produce learning to the other components during the trials with the compound stimulus.

The third type of learning conforms to the paradigm for *negative transfer* and consists of learning different responses to a component and the compound containing that component. The effect on stimulus selection is relevant to the question of whether the subject shifts to a different functional stimulus in the face of negative transfer (Martin, 1968).

EXPERIMENT I

Method

—Lists and design—The design was a $4 \times 2 \times 2$ factorial with type of component list as the first factor, number of trials of component-list learning as the second, and position in the trigram from which the component letters were taken as the third.

First, all subjects learned a paired-associate *practice list* consisting of six vowel-word pairs (A-FIX, E-TAG, I-COW, O-SUM, U-HER, Y-BUD) for either 3 or 10 trials.

Next, the subjects in the experimental conditions learned a 'component list' with one letter from each of the six trigrams in the compound list as stimuli. The subjects in condition A (familiarization) were presented the six letters for free learning. Those in condition A-B (positive transfer) were presented a paired-associate list in which the six letters were paired with the same digits as were the trigrams that contained those letters in the compound list (see below). The subjects in condition A-C (negative transfer) got a paired-associate list in which the six letters were paired with the numbers 7 through 12.

The subjects got 10 or 3 trials on the component list, depending on the number of trials on their practice list, so that all subjects had received 13 learning trials at the end of this stage of the experiment. The subjects in condition C-D got instructions after 3 or 10 trials on the practice list to continue learning the same list for another 10 or 3 trials. The six letters on the component lists for the subjects in conditions A, A-B, and A-C came from either the first or third position in the compound-list trigrams.

Finally, all subjects learned the same *compound list*, six pairs of items with consonant trigrams as stimuli and digits as responses (BQD-1, ZSG-2, CFP-3, XJM-4, TLN-5, RKH-6). There was no duplication of letters in these six trigrams.

—Procedure—All paired-associate learning was by the study/test method with presentation of the six pairs and of the six stimulus items alternating from study to test cycles. Six different sequences were used for presentation of the pairs on the study cycles, and another six sequences for presentation of the stimulus items on the test cycles. All stimuli were presented on two IEE-80 display units at a 2-sec rate with a 2-sec interval between cycles. The subject was not required to pronounce the stimuli as they appeared, and standard paired-associate instructions were used.

In the component-list learning for condition A, the six letters were presented at a 2-sec rate on the right display unit for the study cycles. On the test cycles, a light in the left display unit flashed at a 2-sec rate, and the subject was instructed to say one of the letters each time the light flashed.

The compound list was presented to all subjects for four study/test trials. The subject was not told of the relationship of the compound list to the previous list nor that recall would be required after the learning trials.

Immediately after compound-list learning, each subject was presented the six letters from either the first position of the trigrams or the third position. Each letter was preceded or followed by two dashes representing the missing letters, and the subject was asked to recall the correct *digit* response from the compound list. It should be noted that the set of letters presented as cues for recall was not the set used as stimuli in the component list for that subject. If the subject learned a component list with letters from the first position of the trigrams as stimuli, the letters from the third position were used as recall cues. If the component letters were from the third position, the recall cues were letters from the first position.

After response recall, the same letters were again presented on the display units, and the subject was asked to recall the *letters* represented by the two dashes. Both response and stimulus recall were subject-paced; each letter was presented only after the subject responded to the previous one.

Latencies were recorded for the recall responses in the hope that they might give some information on the possibility of mediated responses (Postman and Greenbloom, 1967). The latencies were characterized only by extreme variability and will not be reported.

—Subjects—The subjects were college students from introductory psychology courses and were not necessarily naive to paired-associate learning. The 16 conditions were listed in counterbalanced orders and subjects assigned to the conditions in that order as they appeared at the laboratory. In addition to the 192 subjects who completed the experiment (12 in each of the 16 groups), 5 subjects were dropped for malfunctions of experimenter or equipment and 1 for failure to learn.

Results

First are the results on *practice-list learning*. All subjects had either three or ten learning trials on the practice list. The mean numbers

of correct responses on the first three trials ranged from 8.58 to 11.67 for the 16 groups; the overall mean was 10.03. On this, as well as subsequent measures, a 'three-between' (type of component list, number of trials of component-list learning, and position of the component letters in the compound) analysis of variance was computed. None of the sources of variance for the number of correct responses on the first three trials approached significance at the .05 level. Thus there was no indication that the groups of subjects differed in learning ability.

For *component-list learning*, there was no reason to expect the lists to be equal in difficulty. More correct responses were given in conditions C-D and A than conditions A-B or A-C. An analysis of variance on the number of correct responses during learning showed that type of component list [$F(3, 176) = 14.99$], number of trials of component-list learning [$F(1, 176) = 1989.50$], and the interaction of type of list and number of trials [$F(3, 176) = 4.83$] were significant.

Three of the 16 groups of subjects made on errors on the last trial of component-list learning, so an analysis of variance on this measure was not appropriate. The mean numbers of correct responses on the last trial for the 8 groups that received three learning trials ranged from 5.33 to 5.92 with an overall mean of 5.60. The means for the 8 groups that received ten learning trials varied from 5.25 to 6.00 with an overall mean of 5.76. The major point here is that performance was close to perfect after three trials and that the seven additional trials by the ten-trial groups were primarily *overlearning* of the component lists. In those 2 groups that had ten trials on the A-B component list, 19 of the 24 subjects made 2 or fewer errors during the last seven trials.

Table 1 presents the mean numbers of correct responses during the four trials of compound-list learning. The means show no indication of any effects due to number of component-list learning trials or to component position. Compared to the C-D control, the free learning

Table 1. Mean correct responses during compound-list learning, Experiment 1

Component list	Component position 1			Component position 3		
	3 trials	10 trials	Combined	3 trials	10 trials	Combined
C-D	17.67	17.58	17.62	16.00	16.33	16.17
A	16.42	16.58	16.50	16.50	15.92	16.21
A-C	13.83	14.42	14.12	17.17	17.25	17.21
A-B	21.25	22.50	21.88	21.92	18.92	20.42
Total	17.29	17.77	17.53	17.90	17.10	17.50

of condition A seemed to have no effect, and condition A-B produced positive transfer. There was a suggestion of negative transfer in condition A-C when the letters from the first position in the trigrams were used in the component list, but not when the third letters were component stimuli.

These rather reasonable trends were supported by the analysis of variance. Type of component list was significant [$F(3, 176) = 13.56$], and the interaction of type of component list and component position approached significance [$F(3, 176) = 2.56, p < .10$]. The error variance (21.62) from the analysis of variance was used in t tests and a difference of 2.70 was required between the combined groups to be significant at the .05 level. With the third-position components, condition A-B was superior to the other conditions. With the first-position components, condition A-B was superior to the other three conditions, and condition A-C inferior to C-D, but not to A.

The mean numbers of correct responses on the response-recall trials are presented in Table 2 for each of the 16 groups and for the combined groups. In parentheses below the heading for the position in the trigram of the letters used as component stimuli is shown the position in the trigram of the letters used as cues for recall. These two positions are redundant in this experiment, but cue position is more important for recall and simplifies discussion of the recall results. Although there were several small differences between the recall to the first- and third-letter cues (Table 2), none of the interactions approached significance. The letters from the first position in the trigrams were more effective cues for response recall than letters from the third position [$F(1, 176) = 97.01$], which agrees with previous findings (Jenkins, 1963; Lovelace and Blass, 1968; Postman and Greenbloom, 1967). The significant effect of type of component list [$F(3, 176) = 13.64$] was primarily due to poorer recall in condition A-B. Increased number of trials of component-list learning facilitated recall

Table 2. Mean correct responses on recall, Experiment I

Component list	Component position 1 (Cue position 3)			Component position 3 (Cue position 1)		
	3 trials	10 trials	Combined	3 trials	10 trials	Combined
C-D	2.58	2.67	2.62	4.83	5.17	5.00
A	1.83	3.25	2.54	4.75	4.42	4.58
A-C	1.92	2.67	2.29	5.25	4.92	5.08
A-B	1.08	1.67	1.38	2.17	3.42	2.79
Total	1.85	2.56	2.21	4.25	4.48	4.36

[$F(1, 176) = 4.58, p < .05$]. Although most of this main effect was due to the difference in the conditions with first-position components (third-letter cues), the interaction of number of trials and component position was not significant [$F(1, 176) = 1.20$].

Table 3 gives the probabilities of correct recall responses to those letters for which neither of the two missing stimulus letters was recalled correctly. The number of cases for the 16 groups ranged from 32 to 53. The probabilities are rather variable but the pattern of differences in recall due to type of component list and to component position is similar to that in Table 2. Apparently these differences cannot be explained by the possibility of differential mediation via letters from other positions (Postman and Greenbloom, 1967).

Table 3. Conditional probabilities of correct recall when no stimulus letters were recalled, Experiment I

Component list	Component position 1 (Cue position 3)			Component position 3 (Cue position 1)		
	3 trials	10 trials	Combined	3 trials	10 trials	Combined
C-D	.32	.28	.30	.67	.80	.74
A	.21	.47	.32	.68	.69	.68
A-C	.23	.20	.22	.85	.77	.81
A-B	.17	.16	.16	.21	.44	.30
Total	.23	.27	.24	.56	.69	.62

On the other hand, the number of trials of component-list learning did not show the same effect on the conditional probabilities as on the number of correct responses recalled. This difference was due to differential recall of the missing stimulus letters. The only significant sources of variance in the number of correct letters recalled were the main effect of number of trials of component-list learning [$F(1, 176) = 6.77, p < .05$] and the interaction with component position [$F(1, 176) = 9.33$]. The main effect here was entirely due to differences in letter recall with the third letters as cues (first-position components). The mean numbers of correct letters recalled to first-letter cues were 3.69 and 3.52 with three and ten trials of component-list learning. The corresponding means for third-letter cues were 2.67 and 4.75. However, in addition to the lack of a significant triple interaction, the trend in the control condition was the same as in the experimental conditions. This makes the significant effects impossible to interpret, since the number of trials of component-list learning was a dummy variable in the control condition. At any rate, the fact remains that superior response recall with the higher degree of component-list learning was related to superior stimulus-letter recall. This could mean that the

superior response recall was due to mediation via other stimulus letters.

In brief, the free learning of a set of component letters produced no effect on compound-list learning or on recall. It is clear that learning the same response to the component as to the compound (condition A-B) produced positive transfer and reduced recall to the other letters. This is not what would be expected if the selector mechanism is relaxed during perfect performance on a paired-associate list. The effect of condition A-C may be more complex, but there was nothing to suggest that the subjects changed functional stimuli in this negative transfer situation. Even when the subjects were pretrained on the nonpreferred set of letters (third-position), there was no increase in recall to letters from the preferred (first) position. However, the recall to the first-letter cues was rather high; about five out of six correct.

EXPERIMENT II

The second experiment was another attempt to determine if, given the opportunity, subjects change functional stimuli from first- to second-list learning, thereby avoiding negative transfer and retroactive inhibition. Consider the retroactive inhibition paradigm where first-list learning is followed by second-list learning with the same nominal stimuli and different responses. If a subject uses one component as the functional stimulus in the first list and immediately shifts to a different component for second-list learning, this should produce less interference with the cued recall of the first-list responses (see Martin, 1968).

Since most subjects select the first letter from a consonant trigram as the functional stimulus for learning, forcing second-list learning to the first letters should produce more retroactive inhibition than forcing learning to letters from the other two positions. If subjects easily shift functional stimuli, then the use of the trigrams as stimuli for learning in both lists provides an opportunity for minimizing retroactive inhibition. In the present experiment, all subjects learned the same first list with consonant trigrams as stimuli. The interpolated list had either the same trigrams, letters from the first position, letters from the second position, or letters from the third position of the trigrams as stimuli.

Method

—Lists and design—The list for *original* learning consisted of six pairs with consonant trigrams as stimuli (LGB, RZM, DSQ, CJT, KHP, XFN) and adjectives

as responses (rapid, quiet, fancy, single, entire, heavy). Each letter in the trigrams occurred only once, and there were three different pairings of stimuli and responses. The *interpolated* list defined the conditions, and the six stimuli consisted of either the same six trigrams as in original learning (condition A-C), the letters from the first position in the trigrams (condition A_1 -C), the letters from the second position in the trigrams (condition A_2 -C), or the letters from the third position (condition A_3 -C). The responses were adjectives different from those in original learning (certain, only, gentle, narrow, direct, proper), and there were two different pairings of the stimuli and responses in each of the four conditions. Each possible original/interpolated list sequence was used equally often.

—Procedure—All stimuli were presented on a Stowe memory drum. Both original and interpolated learning were by the study/test method, with items presented at a 1-sec rate on the study cycles and a 3-sec rate on the test cycles.

The first list was presented to a criterion of five of six correct, and the second list to a criterion of one perfect trial. There was a 10-min interval between the end of original learning and the first recall. The portion of the interval not required for interpolated learning was filled by having the subject count backward by threes.

The stimuli for the first recall were the *trigrams* from the first list. They were presented for one trial at a 4-sec rate; the subject was given a list of the first-list responses and told to guess if he did not know the correct adjective. Three different orders of the trigrams were presented for recall, and an equal number of subjects in each condition received each of the three orders.

Immediately after recall to the trigrams, the 18 *letters* from the trigrams were presented individually and the subject was again asked to give the first-list responses. The subject kept the list of original-learning responses and was asked to guess if he did not know the correct adjective. Each subject was given three recall trials to the 18 letters at a 4-sec rate. Within each trial the sequence of letters was counterbalanced for position in the trigrams. The letters were in a different order for each recall trial, and an equal number of subjects started the recall trials with each of the three orders.

—Subjects—The subjects were college students from an introductory psychology course. The interpolated-learning conditions (A-C, A_1 -C, A_2 -C, A_3 -C) and the control condition were listed in random order, subject to the restriction that each condition appear equally often, and subjects assigned to conditions in that order as they appeared at the laboratory. In addition to 90 subjects who completed the experiment (18 in each of the five conditions), 4 were dropped for failure to complete the learning sequence.

Results

The mean numbers of trials to criterion for original and interpolated learning are shown in Table 4. The trials to criterion on original learning were not significantly different [$F(4, 85) < 1$], so there is no

evidence that the five groups of subjects differed in learning ability. The trials to criterion on interpolated learning were significantly different [$F(3, 68) = 3.43, p < .05$], but since the interpolated learning was different for each condition, the results cannot be interpreted in terms of amount of negative transfer.

Table 4. Mean number of trials to criterion for original learning (first list) and interpolated learning (second list), Experiment II

Condition	First list	Second list
A-C	7.39	8.39
A ₁ -C	8.39	6.28
A ₂ -C	7.22	4.72
A ₃ -C	7.11	5.50
Control	6.44	

The mean numbers of correct responses during the two *recall* tasks are presented in Table 5. The mean numbers of correct responses to the trigrams did not differ significantly as a function of condition [$F(4, 85) = 2.25$], but it should be noted in Table 5 that the condition with the trigrams as stimuli in both lists (condition A-C) produced fewer correct responses to the trigrams than any other condition.

Table 5. Mean number of correct responses recalled to the trigrams and to the single letters from the trigrams, Experiment II

Condition	Trigrams	Single letters			
		A ₁	A ₂	A ₃	Combined
A-C	3.11	6.89	4.50	4.39	15.78
A ₁ -C	3.72	7.72	4.06	4.50	16.28
A ₂ -C	3.89	7.94	4.22	4.11	16.28
A ₃ -C	4.44	10.00	5.61	4.78	20.39
Control	4.39	10.33	4.72	4.22	19.28

The mean numbers of correct recalls to the single letters were 5.89, 5.69, and 6.02 for trials 1, 2, and 3 respectively, differences that did not approach significance ($F < 1$). The mean numbers of correct responses to the letters on all three trials are presented in Table 5 as a function of condition and letter position. The overall recall to the single letters reflects that to the trigrams; neither condition [$F(4, 85) = 1.54$] nor the interaction of condition and letter position [$F(8, 170) = 1.61$] was significant. Position of the letter in the trigram was the only significant source of variance [$F(2, 170) = 70.44$]. As with the trigram stimuli, recall in condition A-C was at least, and Table 5 further shows that the difference in overall recall to the letters was due pri-

marily to differences in recall to letters from the first position.

As a further check on the stimulus selection, an efficiency score was computed from the recall to the single letters for each subject. Only those items on which the recall response was right on two or more of the three trials were considered correct. The efficiency score is the percentage of the total correct recall responses that were 'different' responses. For example, if the three letters from trigram LGB were presented singly and the recall responses to L and B were correct but the response to G was incorrect, that would be two correct recall responses but only one different (in correct recall) response. The number of different sets of three letters from the trigrams that elicited one or more correct responses was divided by the total number of correct responses and the result multiplied by 100. The mean efficiency scores were 81, 81, 80, 72, and 76 for conditions A-C, A₁-C, A₂-C, A₃-C, and the control condition respectively. These differences were not significant ($F < 1$), but the trend for higher efficiency in the experimental conditions suggests that interpolated learning may differentially affect the recall of items with only one versus more than one component as an effective recall cue.

In sum, although the amount of retroactive inhibition was rather small, this should not be allowed to obscure the fact that recall in condition A-C was less than in any other condition. There was again no evidence to suggest that subjects with the same compound stimulus items in the two lists shifted functional components from first- to second-list learning.

EXPERIMENT III

This experiment continued investigation of the effect of the positive transfer involved in learning the same response to the component and to the compound containing the component. Since in Experiment I, only four compound-list learning trials were given to the subjects, it seemed possible that additional trials might reveal a relaxation of the selector mechanism. In addition, the learning to the component stimuli in Experiment III was to a criterion, to ensure that all subjects gave the correct responses to the set of components before learning with the compounds as stimuli.

Method

—Lists and design—The four conditions were defined by the component list used for first-list learning and by the number of trials on the compound

list. The *component* list had either the six letters from the first position of the trigram (condition A₁-B) or the six letters from the third position (condition A₃-B) as stimuli; the responses were the same as those to the corresponding trigrams. After learning of the component list, the compound list was presented for either four or eight learning trials. The *compound* list (A-B) consisted of the same six trigram-digit pairs used in Experiment I.

—Subjects—The subjects were college students and were not necessarily naive to paired-associate learning. Four other conditions from another experiment were run at the same time as the four conditions in this experiment. The conditions were listed in counterbalanced orders and subjects assigned to conditions in that order as they appeared at the laboratory. In addition to the 72 subjects (18 in each of four groups) who completed the experiment, 2 were dropped for failure to learn the first list in 50 trials and 1 because of equipment failure.

—Procedure—Both lists were presented on a Stowe memory drum by the study/test method at a 2-sec rate and a 2-sec intercycle interval. The component list (either A₁-B or A₃-B) was learned to a criterion of two successive perfect trials. Immediately after first-list learning, the trigram-digit list (A-B) was presented for either four or eight learning trials.

After compound-list learning, the 18 letters from the trigrams were presented individually for three recall trials. The position of the letters in the trigrams was counterbalanced within trials, and a different order of the 18 letters was used for each trial. An equal number of subjects within each condition started recall with each of the three orders. Single letters were printed on cards, and the subject was required to turn the cards up one at a time and attempt to give the *digit* that had been paired with the letter during learning. If he did not know the correct response, he was required to guess one of the digits 1 through 6. All responses were recorded by the experimenter.

Results

The mean numbers of trials required to learn the component list to the criterion of two successive perfect trials were 9.67, 7.72, 9.00, and 9.17 for conditions A₁-B with four trials, A₁-B with eight, A₃-B with four, and A₃-B with eight trials respectively. These differences did not approach significance ($F_s < 1$).

The mean numbers of errors during learning of the compound list were 2.28, 2.72, 6.17, and 9.11 for conditions A₁-B with four trials, A₁-B with eight, A₃-B with four, and A₃-B with eight trials respectively. It should be noted that the major difference in number of errors was due to the position of the set of component letters in the trigrams. There were very few errors to the trigrams if component-list learning had been to letters from the first position, and considerably more if it had been to letters from the third position. This difference occurred

primarily during the first four trials of compound-list (A-B) learning. Examination of the distribution of error scores showed that the differences in conditions were almost entirely due to the subjects who made ten or more errors during A-B learning. Only 1 subject from the 36 in condition A₁-B made ten or more errors, while 10 subjects from condition A₃-B made ten or more errors.

The mean numbers of correct *recall* responses were 8.65, 8.39, and 8.61 for trials 1, 2, and 3 respectively; these differences did not approach significance [$F(2, 136) < 1$]. The mean correct responses during the three recall trials are presented in Table 6. Although eight

Table 6. Mean number of correct responses recalled to the letters from each position in the trigrams, Experiment III

Condition	Number of trials	Cue position			Σ
		1	2	3	
A ₁ -B	4	15.94	4.72	4.44	25.11
	8	16.56	6.00	5.61	28.17
A-B	4	8.33	3.56	11.89	23.78
	8	7.78	3.94	13.83	25.56

trials of compound-list learning produced a few more mean correct recall responses than four trials (26.86 and 24.44), the difference was not significant [$F(1, 68) = 2.05, p > .10$]. The only sources of variance that approached significance were trigram position of the letters used as cues for recall [$F(2, 136) = 80.53$] and the interaction of position of recall letters and position of component-list letters [$F(2, 136) = 89.83$]. More correct recall responses were given to letters from the first- and third-position in the trigrams, with the interaction produced by a strong tendency to recall more correct responses to the set of letters used as stimuli in the component list. When adjusted for number of recall trials, the recall to letters from the third position for condition A₁-B and to letters from the first position for condition A₃-B were very similar to the recall for condition A-B in Experiment I (see Table 2).

The mean numbers of correct recalls in Table 6 show that the significant interaction of component list and position of the recall letters was produced primarily by the subjects' better recall to the letters used in the component list. To see if the distributions of correct recall responses to the sets of letters otherwise differed in the two conditions, another ANOVA was computed with the means in Table 6 rearranged so that in each component-list condition the mean recall was ordered the same; that is, the data for cue positions 1 and 3, 2 and 1, and 3 and

2 for conditions A_1 -B and A_3 -B were placed in the same columns. Of course the effect of columns increased, but the interaction remained significant [$F(2, 136) = 13.20$]. This means that learning component lists with letters from the first versus third positions had different effects on learning to the letters in the trigrams—effects apart from the strong tendency to select the set of letters used in the component list. Subjects who learned the A_1 -B component list learned about the same amount to letters in the second and third positions. Subjects who learned the A_3 -B component list learned more to letters in the first position than to letters in the second position.

It would seem that any learning to the trigram letters that were not in the component list would indicate additional learning and that subjects would give the correct response both to the letter from the component list and to letters to which the response was learned during the four or eight trials of the second list. However, the recall (Table 6) to the set of letters used in the component lists was appreciably below the maximum of 18 (for condition A_1 -B the highest was 16.56; for condition A_3 -B, 13.83), and the amount of recall to different letters from the same trigram cannot be determined from the means.

However, efficiency scores were computed for each subject as in Experiment II. The mean scores were 72, 68, 72, and 70 for conditions A_1 -B with four trials, A_1 -B with eight, A_3 -B with four, and A_3 -B with eight trials respectively. Although the eight trials of compound-list learning produced slightly lower efficiency scores than four trials, the difference did not approach significance [$F(1, 68) < 1$]. This means that any additional learning during the eight trials of compound-list learning did not differentially affect the correct recall responses to the three letters from a trigram. Thus, there was again no evidence that the selector mechanism relaxed during A-B learning.

DISCUSSION

The use of a set of components as stimuli in the free-learning task had no detectable effect either on learning to the stimulus compounds or on response recall. Thus, experimental familiarity, or prior exposure to components, did not change cue selection. However, it should be noted that subjects have a strong tendency to select the first letter from consonant trigrams as the functional stimulus and that the stimuli used in this study made such a position rule of selection readily available. If there had been no simple rule available for component selection (if the stimuli had been consonant trigrams in which the order of the

three letters changed on each trial, for example), prior familiarization with one set of components might well have become the basis for selection during learning to the compounds.

The results of the studies just reported are consistent with the assumption that subjects do not change functional stimuli because of the negative transfer produced by learning different responses to the same nominal stimuli. In Experiment I there were no differences in recall to the letters as a function of prior learning of a different response to one component of a compound. In Experiment II the amount of retroactive inhibition was greatest in the condition where both original- and interpolated-learning stimuli were trigrams—in the condition which *permitted* subjects to change functional stimuli from first- to second-list learning and thus minimize retroactive inhibition. Although there is evidence that instructions and prior experience can change the functional stimuli selected for second-list learning (Houston, 1967; Richardson and Chisholm, 1969; Schneider and Houston, 1968; 1969), there is no evidence that selection shifts simply because of negative transfer (Goggin and Martin, 1970; Houston, 1967; Martin, 1968; Martin and Carey, 1971; Weaver, 1969; Wichawut and Martin, 1970; Williams and Underwood, 1970).

Learning a response to a component produced large amounts of positive transfer to learning the same response to a compound containing the component. In Experiment I this positive transfer situation produced less learning to a nonselected component of the compound than that found in a control group. In Experiment III there was no indication that four additional trials of compound learning produced more additional learning to nonselected than to selected components. The components selected from the compound tended to be those used in first-list learning, and overlearning did not produce any relaxation of the selector mechanism. This is not consistent with the results of James and Greeno (1967) and Wichawut and Martin (1970). The problem is one of specifying which conditions do, and which do not produce learning to nonselected components during overlearning.

Notes

This research was supported by National Science Foundation Grants GB5876 and GB17614. Alan Barnett, Elizabeth A. Jones, Kathryn A. Kent, Joseph Platt, and Leonard R. Hersh assisted in the collection of the data.

References

- Goggin, J., and Martin, E. 1970. Forced stimulus encoding and retroactive interference. *Journal of Experimental Psychology* 84:131-136.
- Houston, J. P. 1967. Stimulus selection as influenced by degrees of learning, attention, prior associations, and experience with the stimulus components. *Journal of Experimental Psychology* 73:509-516.
- James, C. T., and Greeno, J. G. 1967. Stimulus selection at different stages of paired-associate learning. *Journal of Experimental Psychology* 74:75-83.
- Jenkins, J. J. 1963. Stimulus "fractionation" in paired-associate learning. *Psychological Reports* 13:409-410.
- Lovelace, E. A., and Blass, E. M. 1968. Utilization of stimulus elements in paired-associate learning. *Journal of Experimental Psychology* 76:596-600.
- Martin, E. 1968. Stimulus meaningfulness and paired-associate transfer: An encoding variability hypothesis. *Psychological Review* 75:421-441.
- Martin, E., and Carey, S. T. 1971. Retroaction, recovery, and stimulus meaningfulness in the A-B, A-Br paradigm. *American Journal of Psychology* 84:123-133.
- Postman, L., and Greenbloom, R. 1967. Conditions of cue selection in the acquisition of paired-associate lists. *Journal of Experimental Psychology* 73:91-100.
- Richardson, J., and Chisholm, D. C. 1969. Transfer of cue selection based on letter position. *Journal of Experimental Psychology* 80:299-303.
- Schneider, N. G., and Houston, J. P. 1968. Stimulus selection and retroactive inhibition. *Journal of Experimental Psychology* 77:166-167.
- Schneider, N. G., and Houston, J. P. 1969. Retroactive inhibition, cue selection, and degree of learning. *American Journal of Psychology* 82:276-279.
- Underwood, B. J., Ham, M., and Ekstrand, B. 1962. Cue selection in paired-associate learning. *Journal of Experimental Psychology* 64:405-409.
- Weaver, G. E. 1969. Stimulus encoding as a determinant of retroactive inhibition. *Journal of Verbal Learning and Verbal Behavior* 8:807-814.
- Wichawut, C., and Martin, E. 1970. Selective stimulus encoding and overlearning in paired-associate learning. *Journal of Experimental Psychology* 85:383-388.
- Williams, R. F., and Underwood, B. J. 1970. Encoding variability: Tests of the Martin hypothesis. *Journal of Experimental Psychology* 86:317-324.

continued from page 498

- Ginsburg, Herbert. *The Myth of the Deprived Child: Poor Children's Intellect and Education*. Englewood Cliffs, N.J.: Prentice-Hall, 1972. Pp. xvi, 252. \$6.95; paperback, \$3.95.
- Gordon, David Cole. *Self-love*. Baltimore: Penguin Books, 1972. Pp. 96. Paperback, \$1.00.
- Gordon, David Cole. *Overcoming the Fear of Death*. Baltimore: Penguin Books, 1972. Pp. 115. Paperback, \$1.00.
- Gorney, Roderic. *The Human Agenda: The Evolution of Human Values: Man's Way to Survival*. New York: Simon and Schuster, 1972. Pp. 698. \$12.95.
- Gottschalk, L. A.; Knapp, P. H.; Reiser, M. F.; Sapria, J. D.; and Shapiro, A. P., eds. *Psychosomatic Classics*. Basel: S. Karger, 1972. Pp. viii, 252. \$8.50.
- Greenfield, Josh. *A Child Called Noah: A Family Journey*. New York: Holt, Rinehart and Winston, 1972. Pp. 193. \$5.95.
- Grene, Marjorie, ed. *Interpretations of Life and Mind: Essays around the Problem of Reduction*. New York: Humanities Press, 1971. Pp. xvi, 152. \$7.50.
- Gross, Alan E.; Collins, Barry E.; and Bryan, James H. *An Introduction to Research in Social Psychology*. New York: Wiley, 1972. Pp. ix, 447.
- Grothan, Martin. *The Voice of the Symbol*. Los Angeles: Mara Books, 1972. Pp. xv, 224. \$8.00.
- Hafen, Brent A., and Faux, Eugene J., eds. *Self-destructive Behavior: A National Crisis*. Minneapolis: Burgess Publishing, 1972. Pp. ix, 330. Paperback, \$4.25.
- Hansburg, Henry G. *Adolescent Separation Anxiety*. Springfield, Ill.: Charles C Thomas, 1972. Pp. xv, 190. \$15.75.
- Harre, R., and Secord, P. S. *The Explanation of Social Behavior*. Totowa, N.J.: Rowman and Littlefield, 1972. Pp. vi, 327. \$13.50.
- Harris, Grace M. *Language for the Preschool Deaf Child*. 3d ed. New York: Grune & Stratton, 1971. Pp. xi, 346. \$9.75.
- Harrower, Molly. *The Therapy of Poetry*. Springfield, Ill.: Charles C Thomas, 1972. Pp. ix, 117. \$4.75.
- Hart, Harold H., ed. *The Complete Immortalia*. New York: Hart Publishing, 1972. Pp. 475. \$12.50.
- Hart, Harold H., introd. *Marriage: For and against*. New York: Hart Publishing, 1972. Pp. 285. \$7.50.
- Hinde, Robert A., ed. *Non-Verbal Communication*. New York: Cambridge University Press, 1972. Pp. xiii, 443. \$17.50.
- Hollander, Edwin P., and Hunt, Raymond G. *Classic Contributions to Social Psychology*. New York: Oxford University Press, 1972. Pp. x, 419. Paperback, \$4.95.
- Hsu, Francis L. K. *Psychological Anthropology*. Cambridge, Mass.: Schenkman Publishing, 1972. Pp. x, 623.
- Hunt, William, ed. *Human Behavior and Its Control*. Cambridge, Mass.: Schenkman Publishing, 1971. Pp. 157.
- Isaacs, Susan. *Social Development in Young Children*. New York: Schocken Books, 1972. Pp. xii, 480. Paperback, \$4.59.

continued on page 548

BACKWARD MASKING: STIMULUS ENERGY AND POSITIONAL UNCERTAINTY

Terry J. Spencer

Kent State University

Backward masking functions were obtained for recognition of single letters off or on foveal center or of one of 12 letters in a circular display; stimulus energy was high (no-mask performance above 95% for the single letters) or low. Stimulus energy, but not foveal location, affected masking ranges for the single letters, yet the same 4:1 energy ratio did not produce differential masking ranges or no-mask performance for the 12-letter displays.

In a discussion of tachistoscopic methodology, Eriksen and Spencer (1969) pointed out that it is often desirable to present visual displays arranged so that all items fall on approximately equally sensitive foveal areas. Displays arranged in a circular manner and centered on the point of fixation tend to satisfy this requirement.

But are results obtained this way generalizable to visual tasks that make more extensive use of the center of the fovea for processing information? Although the sensitivity of the visual system to a letter briefly presented off the foveal center can, by manipulating the intensity or duration of stimulus energy, be equated to that of a letter briefly presented in the center of the fovea, conclusions about certain aspects of the encoding process may not be equally applicable to both locations.

For example, Neisser (1967) suggests that a visual process, such as letter recognition, may involve a sequence of hierarchical stages. At an early stage, preattentive analyzers must first locate and then direct focal attention to a target element for identification. A letter that is presented off the foveal center with positional uncertainty might then require more processing time if locating and directing focal attention to it takes significantly more time than for a letter that is presented on the foveal center with no positional uncertainty. One way to measure this difference would be to compare backward-masking functions by use of a pattern mask designed to control processing time from iconic storage (Spencer 1969; Spencer and Shuntich, 1970). With this procedure, the masking range for a single 'off-center' letter should be greater than that for an 'on-center' one (equated for sensi-

tivity) if it does indeed take longer to process the off-center letter.

Equally important is the effect of varying the energy of the test stimulus on the backward-masking range for both single-letter and circular, 12-letter displays. Previous research (Spencer, 1971) suggests that varying the stimulus energy has little if any effect on masking range when the no-mask, control, performance is 90% or less. However, a stimulus energy sufficient for a no-mask performance of 95% or more may yield a shorter masking range. Uttal (1969a, 1970) has obtained backward-masking ranges for a single-letter stimulus of only 30 to 40 msec as compared to 100 msec or more by others (see, for example Spencer, 1969, or Spencer and Shuntich, 1970). This may have been a result of his using very high (95%) levels of no-mask performance.

Experiment I obtained backward-masking functions for a single-letter display either off or on foveal center and for a 12-letter circular display, each display being presented at two levels of input energy. The purpose was to compare these masking ranges for evidence about the effects of both foveal location and stimulus energy on information processing from iconic storage.

EXPERIMENT I

Method

—Subjects—Four students (two female) with normal or corrected-to-normal vision were recruited from the graduate and undergraduate population at Kent State University and paid for their services. All were given extensive practice in all experimental conditions before the experiment began.

—Apparatus and stimuli—Stimuli were presented in a three-field Scientific Prototype GB tachistoscope. For the 12-letter condition and for the single-letter condition in which the letter was presented off the foveal center, the test stimulus was presented in field 1, the indicator in field 2, and the pattern mask in field 3, all fields being frontlighted. For the single-letter condition in which the letter was presented on the foveal center, no indicator was present, and a foveal mask was presented in field 2. A small and very dim backlighted fixation cross subtending a visual angle of .2 deg was continuously present in the center of field 3 for all conditions.

The stimuli were the letters A, O, V, T (black Paratipe 11316), each represented three times on each of 12 white cards for the condition with the 12-letter displays. The 12 letters were arranged in a circular pattern corresponding to the positions of a clock and were viewed at a distance of 50 in. The diameter of the circle subtended a visual angle of 2.0 deg and the letters a visual angle of .2 deg. The letters were assigned at random to each card, with the restrictions that the same letter never be adjacent to itself and that each occupy all 12 clock positions equally often over the 12 cards.

For the condition with *single off-center letters*, there were 48 single-letter stimulus cards representing all possible combinations of clock positions and letter types, each positioned on an imaginary circle of the same diameter as the 12-letter display. This set was randomly split into two smaller sets of 24 cards, with the restriction that each letter and each position be equally represented. There were also 12 indicator cards, one for each clock position of the test stimulus, simultaneously presented with the 12-letter displays as well as with these displays of single off-center letters. Each indicator was bar-shaped (.12 by .88 deg), white against a flat black background, with the close edge located .47 deg from the outside edge of the test letters.

For the condition with *single on-center letters*, each of the four letters was placed in the center of a separate card. This was repeated six times, to give a set of 24 cards.

A pattern mask was constructed by placing four Paratype Ws back to back and partially overlapping to form a pattern similar to three large adjacent Xs. This pattern was duplicated for each of the clock positions of the test stimulus, all on one card, and was used for masking the 12-letter and the single-letter/off-center presentations. The same pattern was positioned in the center of another card for masking the single-letter/on-center presentations. A flat black card with a circular hole (diameter subtending 2.8 deg) was placed in front of the test-stimulus card, and another in front of the pattern mask. This allowed the circular test field and the pattern mask to be viewed, while providing a dark background for the outside indicator in field 2; the indicator was clearly visible in all exposure conditions. The luminances were set at 25, 20, and 25 mL, for fields 1, 2, and 3 respectively, for the entire experiment.

—Procedure—Backward masking functions were obtained with stimulus onset asynchronies (mask delays) of 0, 25, 50, 75, 100, 125, and 150 msec for single-letter presentations, and 0, 100, 150, 200, 250, and 300 msec for 12-letter presentations. All functions were obtained under two levels of stimulus energy. The *low-energy* condition was defined by placing a .6 neutral density filter (25% transmission) in front of field 1 and varying the stimulus duration for each subject to obtain an approximately 85%–90% performance level for stimulus identification (uncorrected for chance guessing) in the single-letter/on-center, single-letter/off-center, and 12-letter presentations. These durations were determined without a mask (but with a simultaneous indicator in the latter two conditions) and averaged about 1.0 msec for the condition with single on-center letters, 1.8 msec for that with single off-center letters, and 3.0 msec for the 12-letter condition. The *high-energy* condition was defined by simply removing the .6 filter.

Six masking functions were obtained—for single on-center letters, single off-center letters, and 12-letter displays, each at high and low energy—plus six appropriate no-mask (control) conditions. The duration of the masking field was set at 12 msec throughout the experiment, providing a mask far brighter than any of the stimuli. The duration of the indicator field was held constant at 3 msec throughout the experiment.

Data were collected over 36 sessions. There were 12 seven-block (12 trials per block) sessions with the 12-letter display, 6 of those sessions at high energy and 6 at low. Each session contained one block of each of the six

mask delays and one no-mask control, the order of mask delays and controls being partially counterbalanced across sessions and subjects. In addition, there were 24 four-block (24 trials per block) sessions with the single letters, 6 sessions at each combination of energy (high or low) and location (off or on foveal center); half of these sessions contained three of the mask delays and one no-mask control, while the other half contained four of the mask delays. Mask delays and controls were assigned to sessions and subjects with the maximum counterbalancing possible, as was the assignment of session types to presentation order.

Each subject was allowed to dark adapt for 5 min, after which he got 12 warm-up trials without the mask. Before each block he was given 2 practice trials corresponding to the condition to be presented. The order of trials within a block was randomized by shuffling the stimulus cards, and the subject was told his number of correct responses at the end of each block. All trials were self-initiated by the subject after the experimenter gave the ready signal and the subject had the fixation cross in good focus. All responses were recorded by the experimenter.

Results

The percent correct recognition (uncorrected for chance guessing) as a function of stimulus onset asynchrony, or mask delay, for each of the six display conditions is plotted in Figure 1 along with the appropriate no-mask controls. A four-way repeated-measures analysis of variance for the four single-letter functions (foveal location, stimulus energy, mask delay, and subjects) indicated significant main effects for mask delay [$F(7, 21) = 298.20, p < .01$] and stimulus energy [$F(1, 3) = 194.61, p < .01$]. In addition, significant interactions were obtained for stimulus energy and mask delay [$F(7, 21) = 8.86, p < .01$], for stimulus energy and foveal location [$F(1, 3) = 28.09, p < .05$], and for stimulus energy, foveal location, and mask delay [$F(7, 21) = 4.21, p < .05$]. Dunnett's test for comparing all means with a control (Winer, 1962) indicated that both single-letter functions at high energy were not significantly different from their no-mask controls when mask delay was 75 msec or more but that the two single-letter functions at low energy did not asymptote to their controls until mask delay was 100 msec. The effect of increasing stimulus energy for the single-letter conditions was thus both to improve overall average performance and to reduce the masking range by about 25 msec.

The lack of a significant main effect for foveal location suggests that the equating of no-mask (control) performances for off- and on-center letters was maintained under conditions of masking. However, the significant interaction of foveal location and stimulus energy in-

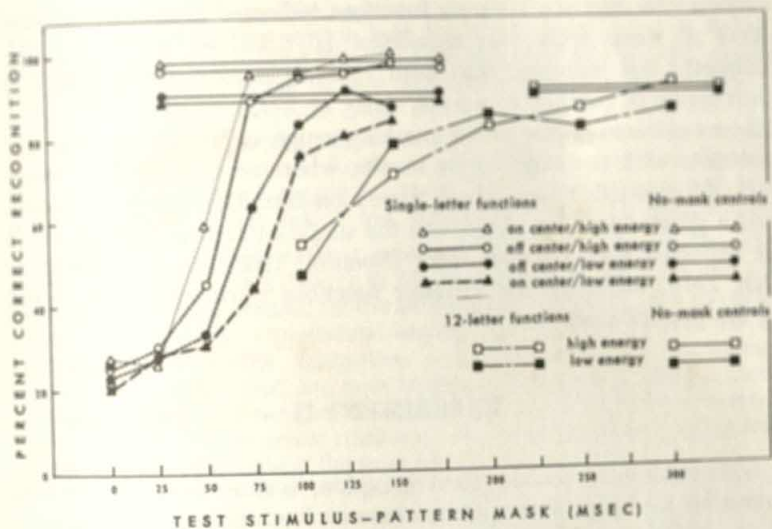


Fig. 1. Percentage of correct identification as a function of stimulus onset asynchrony (mask delay) between test stimulus and pattern mask for the four single-letter functions and the two 12-letter functions of Experiment I (the two 12-letter functions are not extended to the 0-sec mask delay to minimize clutter)

indicated that on the average the two (high and low energy) presentations of single on-center letters were more disparate than the two presentations of single off-center letters. The significant triple interaction indicates this disparity did not obtain for all mask delays.

Tests of the simple interactions of foveal location and stimulus energy were significant ($p < .05$) at mask delays of 50, 75, and 125 msec (at 100 msec it was significant at $p < .10$). Examination of Figure 1 indicates that these simple interactions, except at 50 msec, were due to the subjects' being able, at intermediate mask delays, to identify low-energy letters better when they were presented *off* than on foveal center—a difference that was not evident for the no-mask controls, the condition in which performances for the off- and on-center letters were equated. This was especially surprising because the expectation would be for the on-center condition to 'outperform' the off-center condition if there were any difference, as suggested by the single-letter functions at high energy. Experiment II further investigated this finding.

A three-way analysis of variance (stimulus energy, mask delay, and

subjects) for the two 12-letter functions indicated that only the main effect of mask delay was significant [$F(6, 18) = 149.94$, $p < .01$]. Dunnett's test indicated that both 12-letter functions asymptoted to their no-mask controls at a mask delay of 200 msec. The lack of significant differences for either masking range or overall performance contrasts with the single-letter results, where an equal (4:1) energy ratio differentially affected both these measures. The 12-letter functions had a greater masking range than the single-letter functions, as would be expected from previous work (Spencer, 1969, Spencer and Shuntich, 1970), although the 12-letter functions were somewhat shorter in the present study.

EXPERIMENT II

This experiment was designed to explore in more detail the masking functions for single-letter presentations at low energy. In Experiment I the single on-center letter was superimposed on a very dim back-lighted fixation cross. It was thought that equating off- and on-center performance in the no-mask conditions would control for any possible degradation of the fixation cross. However, it is possible that degradation produced by the fixation cross interacted with stimulus onset asynchrony, or mask delay, perhaps by slowing the rate of contour formation.

Experiment II thus compared the masking range for a single letter presented just slightly *above* foveal center with that of one presented off center, to get evidence on the effect of positional uncertainty independent of any degradation by the fixation cross. This degradation was, in turn, assessed by comparing both the masking ranges and the stimulus duration for above- and on-center presentations. No indicators were used in Experiment II, in contrast to Experiment I, where an indicator was present for off- but not for on-center presentations. Finally, Experiment II used smaller increments (between 50 and 100 msec) of mask delay in order to increase the sensitivity for any potential differences in masking ranges for the three conditions.

Method

—Subjects—Four subjects (two female) with normal or corrected-to-normal vision were recruited from the Kent State University student population and paid for their services. Although three of them had served in previous experiments of a similar nature, none had served in Experiment I.

—Apparatus and stimuli—The apparatus was the same as in Experiment I. The test letters were presented in field 1 and the pattern mask in field 2, both fields being frontlighted. The fixation cross was continuously present in field 3. The stimuli in the *on- and off-center* conditions were the same as in Experiment I. For the *'above-center'* condition, a set of 24 stimulus cards was constructed by duplicating the on-center set but raising the letter slightly so that it appeared above the fixation cross. A pattern mask was also constructed for the above-center presentations. The luminances were the same as in Experiment I, but field 1 was viewed through a 1.0 density filter.

—Procedure—Backward masking functions were obtained with stimulus onset asynchronies (mask delays) of 50, 60, 70, 80, 90, 100, 125, and 150 msec, and a no-mask control, for the above-center, off-center, and on-center conditions. Before the experimental sessions the exposure duration for an approximately 85%–90% recognition accuracy (uncorrected for chance guessing) was established for each subject and display condition. These durations averaged 1.2 msec for the above-center, 1.4 msec for the off-center, and 1.8 msec for the on-center conditions. The duration of the masking field was set at 4 msec throughout the experiment.

Data were collected over 16 sessions. Within each session there were six 24 trial blocks, each of the three conditions being tested at two mask delays (the no-mask control was considered a mask delay), the three blocks of the same delay being run consecutively. The assignment of mask delays to sessions, subjects, and order within sessions, was counterbalanced to the extent possible. There was a total, over subjects, of 288 trials for each combination of mask delay and condition. Except that there was an average of only seven practice sessions, the procedures for warm-up, feedback, trial randomization, dark adapting, and response recording were the same as in Experiment I.

Results

The masking functions for the three (above-, off-, on-center) conditions are plotted in Figure 2. A three-way analysis of variance (foveal location, mask delay, and subjects) indicated that only the main effect of mask delay was significant [$F(7, 21) = 79.92, p < .01$]. Dunnett's test indicated that all three functions asymptoted to their no-mask controls at a mask delay of 100 msec. It is evident that equating the control performances resulted in statistically equivalent masking functions for the three experimental presentations.

A comparison of the average stimulus durations indicated that the above-center condition at 1.2 msec required slightly less energy than the off-center condition at 1.4 msec, with the on-center condition at 1.8 msec requiring the most energy. It is evident that the fixation cross did degrade stimulus input. However, the on-center masking function did not significantly differ from or interact with either the above- or off-center functions, although the small dip at a mask delay of 80

msec was in the expected direction. The results are in agreement with Experiment I in that the off-center function did not require a greater mask delay than either the above- or on-center conditions, again showing no differential masking range as a function of positional uncertainty.

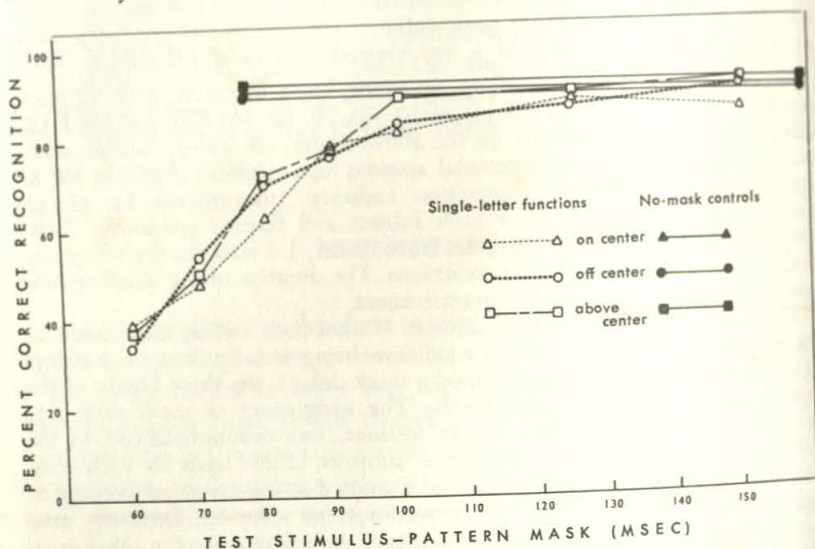


Fig. 2. Percentage of correct identification as a function of stimulus onset asynchrony (mask delay) between test stimulus and pattern mask for the three single-letter functions of Experiment II

DISCUSSION

The expectation that a single letter presented off the foveal center would require a greater mask delay than one presented in the center of the fovea was not confirmed in either Experiment I or II. The subjects were able to locate and process an off-center letter with positional uncertainty on the circumference of an imaginary 2-deg circle in the same 'time' as an on- or above-center letter without positional uncertainty. (Time in this context refers to masking range.) Yet they were not able to locate and process a critical letter as rapidly in the 12-letter condition.

It is evident that the difference in masking ranges for the single- and 12-letter conditions was not simply a function of positional uncertainty, which was equivalent in the 12-letter and the single-letter/

off-center conditions. Rather, the difference would seem to be in the speed or ease of reducing the positional uncertainty in each case. That is, location of the target element in the 12-letter condition required the processing of an indicator. The evidence (Eriksen and Collins, 1969) that it takes approximately 100 to 200 msec to encode an indicator and direct attention to a critical element in a multielement display would thus account nicely for the differences in masking ranges for single letters and 12-letter displays—and Experiment II confirmed that subjects did not require an indicator in conditions with a single off-center letter.

However, the more difficult problem is to account for the lack of a difference in masking ranges for single on- and off-center letters. It was not due to degrading by the fixation cross, since the masking functions for the above- and off-center as well as on-center presentations in Experiment II were essentially identical. Yet it seems unlikely that positional uncertainty has no effect on processing, or that attention can be instantaneously directed to a letter that is off the foveal center. It seems more likely that the present masking procedure was not sensitive to the directing of focal attention in the conditions with a single off-center letter.

If it can be assumed that contour formation, and the representation of the icon at the higher centers, requires a significant time (say, 50 to 75 msec, depending on various factors), focal attention may be directed toward a single off-center letter *before* development of the icon is complete. That is, locating a single letter may require only a preliminary stage of contour formation, with attention being directed during the remainder of the icon development. The masking range in this study was a measure of the time to encode the *identity* of a letter, not simply to locate the *position* of a letter. Thus, attention may have been fully directed to the single letters in both the on-center and off-center conditions by the time the icon was sufficiently developed for identification. More research is needed to determine if there is indeed a differential masking range for the presence or absence, as compared to the identification, of a single letter in the on- versus off-center positions.

The finding of no differences in masking range for off-center versus on- and above-center letters suggests that, with respect to encoding from iconic storage, results obtained with a single letter that is presented on the circumference of an imaginary 2-deg-diameter circle centered on the fovea may be comparable to the results obtained with a single letter that is itself presented centered on the fovea if the two

presentations are equated for sensitivity. However, this may be true only for tasks requiring full identification of a target stimulus if, as argued above, the present experiment was insensitive to temporal requirements for directing focal attention to an off-center location.

The results for the single-letter conditions in Experiment I indicate that the masking range was dependent on the level of no-mask (control) performance. Increasing performance to over 95% significantly reduced the 'apparent' masking range. The term apparent is used because the mask may have been partially reducing the information processed from the high-energy displays at a mask delay of 75 msec, as it did for the low-energy displays. However, this information loss may not have been evident in performance due to the nonlinearity of recognition measures and psychological processes at performance levels near 100% (Eriksen and Collins, 1969; Eriksen and Spencer, 1969). At these levels, relatively large increases or decreases in stimulus quality may result in only small changes in performance. Thus, even with a reduction in stimulus quality at a mask delay of 75 msec for the high-energy condition, enough information may have been processed for a high recognition score. With greater mask delays, more information may have been processed but not reflected in performance, which had already asymptoted at or near the 100% level at a mask delay of 75 msec.

On the other hand, there is also the possibility that the high-energy functions for the single letters asymptoted sooner as a result of faster transmission time to the higher centers. Kahneman (1968, p. 406) cites evidence suggesting that "conduction speed is usually a function of stimulus intensity." A faster conduction speed would be equivalent to increasing mask delay by a comparable amount. Whether this factor was solely or even partially responsible for the different asymptotes could not be determined in the present study. However, Spencer (1971) has found that with even greater variations in stimulus energy than in the present study, but with all conditions maintained at performance levels of less than 90% (by varying task difficulty), there were no differential asymptotes of single-letter masking functions. Similarly, Spencer and Shuntich (1970) found equivalent masking at delays of 150 msec or more by pattern masks of very different energies, which argues against differential transmission times for each mask.

It is, of course, possible that both factors—insensitive high performance levels and faster transmission times—combined to produce the differential asymptotes in the present study. It is interesting in this context to examine Uttal's (1969b, 1970) failure to obtain an algebraic

combination of forward and backward masking. His separately obtained estimates of the maximum forward- and backward-masking ranges may have been 'underestimated' as a result of using displays on which the no-mask controls were at or near 100% performance levels. Thus, masking may have been present at delays of ± 50 msec, for example, but gone undetected in performance. However, the combined effects of forward and backward masking at the same delays (that is, a 100-msec 'hole') may have been sufficient to reduce performance to a measurable region. It must be recognized that many other factors could also be responsible for Uttal's results, not the least of which is his very different type of display. Nevertheless, it is important that experimenters be cognizant of this possibility when obtaining estimates of masking ranges.

The results for the two 12-letter functions indicate that varying the test-stimulus energy in an approximately 4:1 ratio did not produce significant differences either in no-mask, control, performance or in the masking ranges. Sperling (1960) also reported that stimulus duration did not have a potent effect on partial report from multi-element displays, as did Scharf and Lefton (1970) and Eriksen and Rohrbaugh (1970). Keele and Chase (1967) found only small differences in recognition performance or estimated iconic storage time for their two lower-energy presentations, which also differed by an approximately 4:1 ratio. Only with a much larger increase in stimulus energy, 16:1, did they find a large difference in target identification from a multielement display.

Unexplained is why performance in the single-letter conditions was sensitive to this energy variation but performance in the 12-letter condition was not. It may have been the result of the subjects' having had to locate the indicator before they could direct their focal attention toward the critical letter in the 12-letter condition. If locating such an indicator takes approximately 100-200 msec (Eriksen and Collins, 1969), and if a subject so depends on having an indicator that delaying it 100 msec requires an additional 100-msec delay of a pattern mask (Spencer, 1969), then it seems likely that a subject would not begin processing the critical letter as quickly in a 12-letter as in a single-letter condition. This processing delay for a 12-letter display would be accompanied by a fading of the iconic storage in both high- and low-energy presentations.

If we may then assume that the *rate* of iconic decay is directly proportional to the original icon's energy, the high-energy stimulus would fade to a level of strength very near that of the low-energy

icon by the time a subject directed his focal attention to the critical letter. Admittedly there is little evidence on this assumption of iconic-decay rate, but it appears to be consistent with the existing and relevant data. Additional research is required.

Finally, a tentative generalization that seems to fit much of the outcome of the present study is that if two or more single-letter (or 12-letter) conditions are comparable in their levels of no-mask, control, performance, regardless of their actual energy ratios, equivalent backward-masking ranges will be obtained. Whether this generalization will survive further testing remains to be seen.

Notes

The author wishes to thank Richard Shuntich, Lawrence Hawkes, and Gregory Mattson for their able assistance in the collection and analysis of the data. Requests for reprints should be sent to Terry J. Spencer, Department of Psychology, Kent State University, Kent, Ohio 44242. Revision received for publication August 31, 1971.

References

- Eriksen, C. W., and Collins, J. F. 1969. Temporal course of selective attention. *Journal of Experimental Psychology* 80:254-261.
- Eriksen, C. W., and Rohrbaugh, J. 1970. Visual masking in multielement displays. *Journal of Experimental Psychology* 83:147-154.
- Eriksen, C. W., and Spencer, T. 1969. Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology* 79(part 2):1-16.
- Kahneman, D. 1968. Method, findings, and theory in studies of visual masking. *Psychological Bulletin* 70:404-425.
- Keele, S. W., and Chase, W. G. 1967. Short-term visual storage. *Perception and Psychophysics* 2:383-386.
- Neisser, U. 1967. *Cognitive psychology*. New York: Appleton-Century-Crofts.
- Scharf, B., and Lefton, L. A. 1970. Backward and forward masking as a function of stimulus and task parameters. *Journal of Experimental Psychology* 84:331-338.
- Spencer, T. J. 1969. Some effects of different masking stimuli on iconic storage. *Journal of Experimental Psychology* 81:132-140.
- Spencer, T. J. 1971. Encoding time from iconic storage: A single letter visual display. *Journal of Experimental Psychology* 91:18-24.
- Spencer, T. J., and Shuntich, R. 1970. Evidence for an interruption theory of backward masking. *Journal of Experimental Psychology* 85:198-203.
- Sperling, G. 1960. The information available in brief visual presentations. *Psychological Monographs* 74:1-29.
- Uttal, W. R. 1969a. Masking of alphabetic character recognition by dynamic visual noise (DVN). *Perception and Psychophysics* 6:121-128.

- Uttal, W. R. 1969b. The character in the hole experiment: Interaction of forward and backward masking of alphabetic character recognition by dynamic visual noise (DVN). *Perception and Psychophysics* 6:177-181.
- Uttal, W. R. 1970. Masking of alphabetic character recognition by ultra-high-density dynamic visual noise. *Perception and Psychophysics* 7: 19-22.
- Winer, B. J. 1962. *Statistical principles in experimental design*. New York: McGraw-Hill.

continued from page 534

- Ivey, Allen E. *Microcounseling: Innovations in Interviewing Training*. Springfield, Ill.: Charles C Thomas, 1972. Pp. xxi, 206. \$9.75.
- James, Charity. *Young Lives at Stake*. New York: Agathon Press, 1972. Pp. xx, 258. \$6.95.
- Johnson, Ronald C.; Dokecki, Paul R.; and Mowrer, O. Hobart, eds. *Conscience, Contract, and Social Reality: Theory and Research in Behavioral Science*. New York: Holt, Rinehart and Winston, 1972. Pp. xiv, 460.
- Johnston, William M. *The Austrian Mind: An Intellectual and Social History*. Berkeley and Los Angeles: University of California Press, 1972. Pp. xv, 515. \$17.50.
- Jones, N. Blurton, ed. *Ethological Studies of Child Behaviour*. New York: Cambridge University Press, 1972. Pp. x, 400. \$19.50.
- Jordan, Nehemiah. *Themes in Speculative Psychology*. New York: Harper & Row, 1972. Pp. 259. \$6.75; paperback, \$2.50.
- Kagan, Jerome, and Havemann, Ernest. *Psychology: An Introduction*. 2d ed. New York: Harcourt Brace Jovanovich, 1972. Pp. xii, 627. \$10.50.
- Kahn, Herman, and Bruce-Briggs, B. *Things to Come: Thinking about the 70's and 80's*. New York: Macmillan, 1972. Pp. 262. \$6.95.
- Kahn, Theodore C. *An Introduction to Hominology: The Study of the Whole Man*. Springfield, Ill.: Charles C Thomas, 1971. Pp. xvii, 384. \$12.50.
- Kane, J. E., ed. *Psychological Aspects of Physical Education and Sport*. Boston: Routledge & Kegan Paul, 1972. Pp. x, 237. \$8.25.
- Kanner, Leo. *Child Psychiatry*. 4th ed. Springfield, Ill.: Charles C Thomas, 1972. Pp. xxiv, 735. \$16.50.
- Kaplan, Arnold, ed. *Genetic Factors in "Schizophrenia."* Springfield, Ill.: Charles C Thomas, 1972. Pp. xxvi, 698. \$43.00.
- Keats, J. A. *An Introduction to Quantitative Psychology*. New York: Wiley, 1972. Pp. xii, 126. Paperback, \$3.95.
- Kennedy, Pat. *Child Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1972. Pp. xi, 452. Paperback, \$5.95.
- Kissin, Benjamin, and Begleiter, Henri, eds. *The Biology of Alcoholism*. Vol. 1, *Biochemistry*. Vol. 2, *Physiology and Behavior*. New York: Plenum Publishing, 1972. Pp. xxviii, 630; xxx, 552. \$32.50, \$37.50.
- Kling, J. W., and Riggs, Lorrin A., eds. *Woodworth and Schlosberg's Experimental Psychology*. 3d ed. New York: Holt, Rinehart and Winston, 1971. Pp. xiv, 1,279.
- Kohler, Wolfgang. *The Task of Gestalt Psychology*. Princeton, N.J.: Princeton University Press, 1972. Pp. vii, 164. Paperback, \$2.95.
- Kraines, Samuel Henry, and Thetford, Eloise S. *Help for the Depressed*. Springfield, Ill.: Charles C Thomas, 1972. Pp. viii, 260. \$8.25.
- Kramer, John, ed. *North American Suburbs: Politics, Diversity, and Change*. Berkeley, Calif.: The Glendessary Press, 1972. Pp. xxi, 330. \$7.95; paperback, \$3.95.
- Krawiec, T. S., ed. *The Psychologists*. Vol. 1. New York: Oxford University Press, 1972. Pp. viii, 376. \$8.00; paperback, \$3.95.

continued on page 568

MARBE'S LAW, STIMULUS EXPECTANCY, AND RESPONSE COMPETITION

Donald W. Stilson, Kjell Haaseth,
Charles R. Walsmith, Robert J. Schneider, and Michael Rogers
*University of Colorado School of Medicine, Gaustad Hospital,
and University of Colorado School of Medicine*

The hypothesis was that word-association response latency is an increasing linear function of overt response competition (measured by conditional response uncertainty) and of stimulus unexpectedness (measured by negative log stimulus probability). Both relationships were monotonic-increasing, as predicted, but neither was linear. Rather, response speed (reciprocal latency) decreased linearly with overt response competition and with stimulus unexpectedness. A more general, more precise form of Marbe's law is proposed, and a neurophysiologically oriented model of word-association behavior is sketched.

It has been amply demonstrated that word-association response latency depends on response-word frequency. Thumb and Marbe (1901) were the first to report a decrease in latency with increasing response frequency, a result verified since by Wreschner (1907), Menzerath (1908), Cason and Cason (1925), and Schlosberg and Heinemann (1950). This relationship is sometimes referred to as the Thumb-Marbe law, or simply as Marbe's law.

Baker and Elliott (1948) found shorter latencies for the same stimulus-response pairs when subjects were instructed to restrict their responses to opposite meanings than when no such restriction was imposed. This suggests that latency increases with response competition. Moreover, it seems intuitively plausible that a high-probability response occurs with less competition from other 'response tendencies' than does a low-probability one—which may account for Marbe's law. More generally, a stimulus word that evokes a large number of equiprobable responses provides evidence of more response competition than one with a single response that occurs with probability 1. This in turn suggests that conditional response uncertainty may reflect word-association response competition. If $P(R_i|S_j)$ is the probability that response i will be given when stimulus j is presented, and r_j is the number of distinct responses with nonzero probability, then the con-

ditional response uncertainty associated with stimulus j is

$$H(R|S_j) = -K \sum_{i=1}^{r_j} P(R_i|S_j) \log_2 P(R_i|S_j). \quad [1]$$

With K set equal to the mean strength of competing responses, Equation 1 becomes the measure of conflict proposed by Berlyne (1957; 1960, pp. 34-37). In subsequent discussion, we make the simplifying assumption that $K = 1$.

That such a measure may indeed be related to word-association response latency is shown by Laffal's (1955) report of a linear correlation of .77 between an estimate of $H(R|S_j)$ and the frequency of latencies above 2.5 sec. Laffal also noted a correlation of .78 between the number of distinct responses (an estimate of r_j) and the frequency of long latencies. We calculated similar correlations (.70, .61, and .63) between $\hat{r}_j$ and mean latency for three sets of word-association data published by Crane (1915). Thus, it appears that $H(R|S_j)$ and r_j each covary with response latency—and both relationships are consistent with the suggestion that word-association response latency increases with increasing response competition. Accordingly, Hypothesis I of the present study was that word-association response latency is a linear, increasing function of response competition measured by $H(R|S_j)$.

Although there have been several demonstrations of covariation between response frequency and latency, we found only one report of a relationship between stimulus-word probability and word-association response latency. Hall and Ungelow (1957) found longer word-association response latencies to stimulus words of lower frequency in the Thorndike-Lorge (1944) counts than to those of higher frequency. This suggests that response latency increases with the unexpectedness of the stimulus word. As a measure of unexpectedness, we used the negative log of stimulus-word probability, which Attneave (1959, p. 6) calls surprisal. Thus, our Hypothesis II was that response latency is a linear, increasing function of stimulus-word surprisal.

METHOD

—Subjects—The English-speaking subjects were 26 employees of the University of Colorado Medical Center; 32 psychiatric patients newly admitted to the Colorado Psychiatric Hospital and with a subsequent discharge diagnosis other than psychosis; and 14 psychotic patients newly admitted to the same hospital, 10 of them with a discharge diagnosis of schizophrenia and

4 with one of psychotic depression. None of the patients acknowledged taking drugs in the last few weeks before their hospitalization or received any drugs from the time of admission to the time of their participation in the experiment.

The Norwegian-speaking subjects were 35 young men inducted into the Norwegian navy six to eight weeks before testing, and another 35 young men recruited for the navy and subsequently rejected due to a medical board's decision that they suffered from neurosis or psychic instability. Age, education, and sex-distribution statistics for all the subjects are shown in Table 1.

Table 1. Characteristics of the subject samples

	Age		Years of education		Percent males
	Mean	SD	Mean	SD	
English-speaking subjects					
Employees ($N = 26$)	29.2	9.3	13.2	2.0	47
Nonpsychotics ($N = 32$)	32.8	8.9	11.2	2.1	31
Psychotics ($N = 14$)	32.2	9.2	11.6	4.0	64
Schizophrenics ($N = 10$)	28.2	7.4	12.3	3.5	80
Depressed ($N = 4$)	42.5	4.2	10.0	5.5	25
Norwegian-speaking subjects					
Normal recruits ($N = 35$)	21.2	1.4	8.7	1.1	100
Rejected recruits ($N = 35$)	21.7	1.4	8.8	1.0	100

—Stimuli—For the English-speaking subjects, 50 stimulus words were chosen from those in the study by Palermo and Jenkins (1964). Words for the Norwegian-speaking subjects were translations of 31 words from the Kent and Rosanoff (1910) list. (There were also 4 specially selected words that do not concern us in the study now being reported.) The stimulus words appear in Table 2, which also gives the frequency of occurrence of each English word in the Lorge magazine count (Thorndike and Lorge, 1944) and the primary English equivalent of each Norwegian word.

—Procedure—The subjects were studied individually. Instructions and stimulus words were presented to the English-speaking subjects by means of a tape recorder. Each subject was instructed to respond with the first single word that occurred to him after presentation of the stimulus word and was asked not to give other grammatical forms of the stimulus word, nor to repeat it. Three practice words (blotter, ink, ashtray) were then presented. If the subject delayed more than a few seconds, the experimenter stopped the tape recorder and encouraged him by saying, "Say the word that 'blotter' [for example] makes you think of."

The instructions read by the experimenter to the Norwegian-speaking subjects were similar, but grammatical variants of the stimulus word were not explicitly excluded and the subject was told that his responses were being timed. The stimulus words were presented orally by the experimenter. Immediately after completing the procedure, each Norwegian subject was

Table 2. Stimulus words and their characteristics; conditional response uncertainty is estimated from the twelfth-grade norms of Palermo and Jenkins (1964) for English and from Haaseth's (1968) norms for Norwegian

English words (English equivalents)	Frequency in Lorge magazine count	Conditional response uncertainty, men's norms	Conditional response uncertainty, women's norms	Norwegian words	Conditional response uncertainty
anger	234	3.87	3.73	sinne	5.94
beautiful	987	4.44	4.49		
black	1,083	3.31	3.07		
butterfly	65	4.92	4.54		
cabbage	114	4.45	4.10		
carpet	165	3.10	3.18		
chair	1,298	3.47	3.09		
citizen	332	4.64	4.49		
cold	1,092	3.52	3.26		
comfort	508	5.10	5.06		
command	301	5.17	5.04		
dark	1,005	2.27	2.23	mørke	3.80
deep	881	4.15	4.26		
eagle	101	2.82	2.53		
earth	448	4.44	4.63		
eating	816 ^a	3.91	3.77		
foot	505	4.13	3.97		
fruit	630	3.87	3.30	jente	4.78
girl	4,058	3.73	2.86		
hand	3,334	4.07	3.98		
hard	1,909	3.43	3.13		
high	1,674	3.21	3.49	hus	5.05
house	4,231	5.11	4.99		

man	7,355	3.03	3.11		
mountain	288	4.33	4.23		
music	660	5.25	4.71		
mutton	22	4.40	4.28		2.31
needle	88	3.20	2.67		
red	1,036	4.82	4.60		
river	455	4.06	4.11		
rough	294	4.73	4.25		
short	887	3.19	3.02		
sickness	62	3.99	3.93	sykdom	5.40
sleep	503	4.48	4.24	søvn	4.48
slow	434	2.81	2.88	jevn	5.04
smooth	346	3.84	3.94		
soft	549	4.15	4.42	soldat	5.52
soldier	259	4.41	4.24		
sour	102	3.63	3.34		
spider	42	3.84	4.34	mage	5.12
stomach	148	4.02	4.34		
sweet	679	3.76	3.84	bord	2.73
table	1,325	2.53	2.02		
trouble	1,180	5.85	6.37		
whistle	211	5.27	5.33		
white	2,663	3.19	3.28		
window	1,564	4.26	4.39	ønske	5.51
wish	1,220	5.08	4.95	kvinne	4.59
woman	2,431	3.26	3.26		
working	2,720 ^b	4.87	5.15	arbeide	5.14
(work)				barn	5.10
(child)				bibel	4.09
(bible)					

Table 2. (Continued)

English words (English equivalents)	Frequency in Lorge magazine count	Conditional response uncertainty, men's norms	Conditional response uncertainty, women's norms	Norwegian words	Conditional response uncertainty
(liquor)				brennevin	5.15
(dream)				drøm	4.71
(square)				firkant	4.75
(delight)				fryd	3.63
(hammer)				hammer	3.77
(ocean)				hav	4.12
(health)				helse	4.89
(lamp)				lampe	1.87
(order)				ordre	5.62
(priest)				prest	5.24
(afraid)				redd	5.39
(justice)				rettferdighet	5.53
(slowly)				sakte	4.37
(stem)				stilk	4.07
(difficulty)				vanskelighet	6.35

^a'Eating' not listed separately. The frequency is for 'eat' in its various verb forms.

^b'Working' not listed separately. The frequency is for 'work' in its various forms.

asked to reproduce his responses in a second presentation of the stimulus words. Response latencies were measured with a stopwatch for all of the subjects.

One psychotic patient was dropped from the first data analysis since she failed to respond to 11 words. Three hospitalized nonpsychotic patients and one employee were all unfamiliar with the same stimulus word (mutton), but their 49 other responses were included.

RESULTS

Hypothesis I

Using Palermo and Jenkins' (1964) sex-appropriate word-association response frequencies for twelfth graders, $H(R|S_j)$ was estimated for each English stimulus word.¹ The same procedure was followed for the Norwegian words using Haaseth's (1968) word-association response frequencies.

Plots of the bivariate distribution over stimulus words of mean latency and $\hat{H}(R|S_j)$ revealed that for all eight groups of subjects the relationship was described quite well by a linear function for $\hat{H}(R|S_j) < 4.5$ bits. Above this point, however, there was strong evidence of nonlinearity. Because of this, we were wary of further analysis of averaged data, and intraindividual analyses were undertaken.

Since, as Equation 1 shows, $H(R|S_j)$ is the expected value of conditional response surprisal, $-\log_2 P(R_i|S_j)$, the relationship between latency and $-\log_2 P(R_i|S_j)$ is the intraindividual analog of the relationship between mean latency and $H(R|S_j)$ reported above. A plot of the bivariate distribution of estimates of conditional response surprisal and latency was examined for each subject.² These provided consistent and unequivocal graphical evidence of nonlinearity. However, the relationship between response speed (reciprocal latency) and conditional response surprisal was approximately linear for most subjects.³

For 141 of the 142 subjects, the linear correlation between speed and conditional response surprisal was negative. The mean correlation varied from $-.31$ to $-.49$ in the eight samples of subjects, male and female (see Table 3), and from $-.46$ to $-.49$ in the three psychiatrically normal samples. Assuming that $P(R_i|S_j)$ is the same for all subjects, the expected speed of response i to stimulus word j for subject k is

$$\mu_{ijk} = A_k [-\log_2 P(R_i|S_j)] + B_k, \quad [2]$$

Table 3. Summary of intraindividual correlations

	Response speed with conditional response surprisal		Response speed with stimulus surprisal		Conditional response surprisal with stimulus surprisal	
	Males	Females	Males	Females	Males	Females
English-speaking samples						
Employees						
N	12	14	12	14	12	14
Mean	-.49	-.46	-.19	-.22	.04	.12
SD	.09	.11	.19	.15	.13	.13
Nonpsychotics						
N	10	22	10	22	10	22
Mean	-.48	-.47	-.21	-.20	.08	.07
SD	.15	.11	.16	.16	.18	.19
Psychotics						
N	9	5	9	5	9	5
Mean	-.40	-.31	-.15	-.12	.03	.03
SD	.18	.16	.15	.18	.18	.13
Norwegian-speaking samples						
Normal recruits						
N	35					
Mean	-.48					
SD	.14					
Rejected recruits						
N	35					
Mean	-.38					
SD	.18					

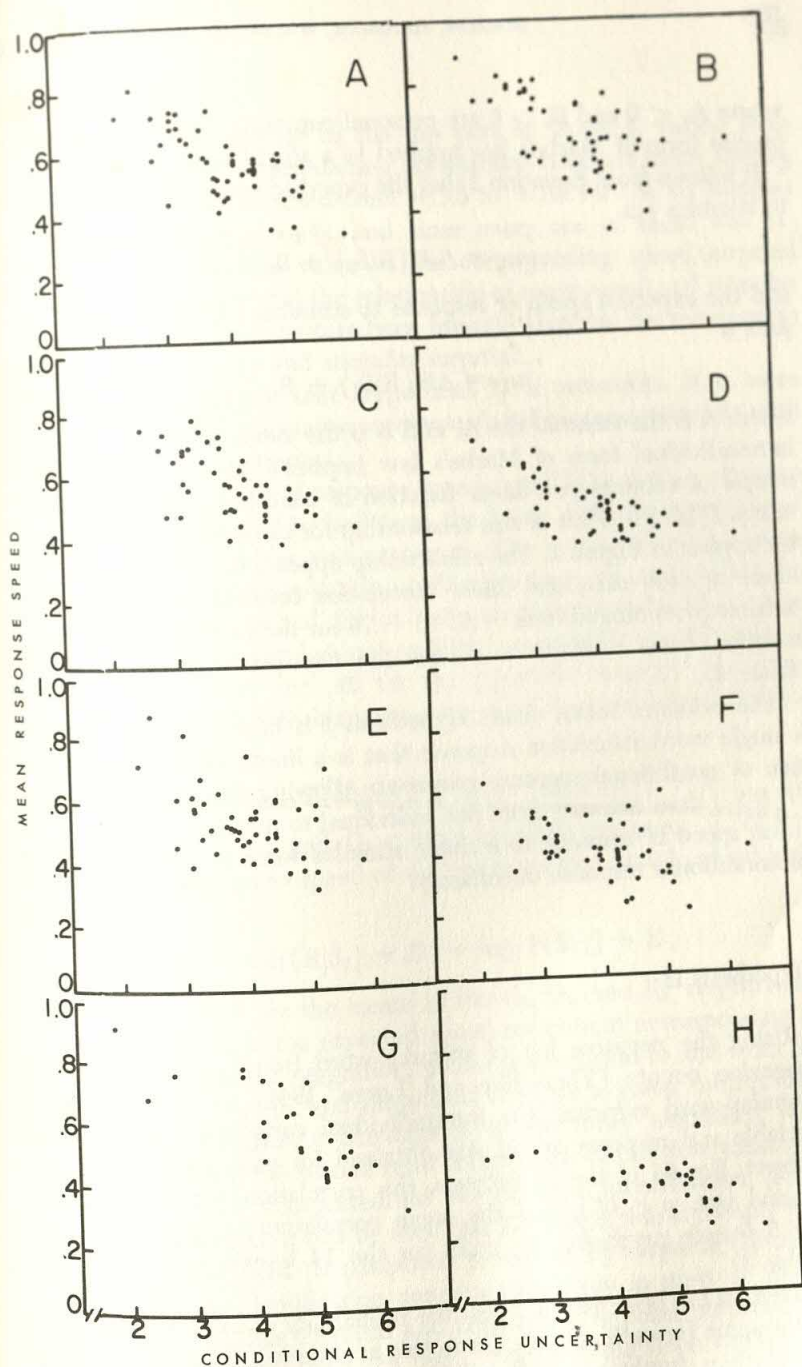


Fig. 1. Plots of the relationship between mean response speed and conditional response uncertainty for (A) male employees, (B) female employees, (C) male nonpsychotics, (D) female nonpsychotics, (E) male psychotics, (F) female psychotics, (G) accepted recruits, and (H) rejected recruits

where $A_k < 0$ and $B_k > 0$ are personal constants. This, then, is a more precise form of Marbe's law applied to a single subject.

It follows from Equation 2 that the expected speed over all responses to stimulus j is

$$\mu_{\cdot jk} = A_k H(R|S_j) + B_k,$$

and the expected speed of response to stimulus j for a sample of subjects is

$$\mu_{\cdot j\cdot} = AH(R|S_j) + B,$$

where A is the mean of the A_k and B is the mean of the B_k . Thus, the intraindividual form of Marbe's law implies that mean speed for a sample of subjects is a linear function of conditional response uncertainty, $H(R|S_j)$. Plots of this relationship for each group of subjects by sex appear in Figure 1. The relationship appears to have been roughly linear in each case, and linear correlation coefficients corresponding to these plots ranged from $-.61$ to $-.78$ for the psychiatrically normal samples, from $-.48$ to $-.49$ for the two psychotic samples (see Table 4).

The evidence seems clear: Hypothesis I is false. It is the *speed* of a single word-association response that is a linear (decreasing) function of conditional response surprisal. Allowing the assumption that $H(R|S_j)$ does not vary from one individual to the next, it follows that mean speed of response to a single stimulus word is a linear function of conditional response uncertainty.

Hypothesis II

Using the negative log of stimulus-word frequency in the Lorge magazine counts (Thorndike and Lorge, 1944) as an estimate of stimulus-word surprisal, the intraindividual correlation between this variable and response speed was obtained for each English-speaking subject. For 61 of the 72 subjects, this correlation was negative (binomial test, $p < .05$), and the mean correlation ranged from $-.12$ for 5 female psychotics to $-.22$ for the 14 female employees (see Table 3).

Here, it has been assumed that the probability of each stimulus word is the same for each subject. Given this assumption, the intraindividual correlations imply a negative linear correlation between the mean (over subjects) of response speed and stimulus-word surprisal. These

correlations, which appear in the top part of Table 4, varied from $-.35$ to $-.41$. The corresponding correlations between mean latency and stimulus surprisal varied from $+.15$ to $+.35$ for the six samples of English-speaking subjects, and since every one of these was exceeded by the absolute value of the corresponding speed/surprisal correlation, it appears that the relationship of mean speed and stimulus surprisal showed less departure from linearity than did the relationship between mean latency and stimulus surprisal.

These results suggest that Hypothesis II is untenable. It is more plausible that word-association response *speed* decreases linearly with increasing stimulus surprisal.

Since the regressions of response speed on $H(R|S)$ and on stimulus surprisal both appeared to be linear, the linear multiple regression of speed on these variables was examined. The results are summarized in Table 5. Multiple correlation coefficients were $.69$ and $.74$ for the two employee samples and $.58$ for both psychotic samples. The means of the analogous intraindividual multiple correlations are $.50$ and $.53$ for the normals, $.35$ and $.46$ for the psychotic samples. Thus, the expected speed with which response i is given by person k to stimulus j can be expressed as

$$\mu_{ijk} = C_k [-\log_2 P(R_i|S_j)] + D_k [-\log_2 P(S_j)] + E_k, \quad [4]$$

where $E_k > 0$, $C_k < 0$, and $D_k < 0$ are personal constants. It follows from this that the mean speed of the responses of a sample of subjects to stimulus j is

$$\mu_{.j.} = CH(R|S_j) + D [-\log_2 P(S_j)] + E,$$

where C , D , and E are the means of the C_k , D_k , and E_k respectively.

The constant E_k is the expected speed component corresponding to the sum of peripheral integration, neural transmission to and from the central nervous system, and movement time for a given subject. The first term in Equation 4 corresponds to the delay associated with response selection; the second term corresponds to the time consumed by the effects of unexpectedness on stimulus identification. From these interpretations, it can be seen that Equations 4 and 5 imply that the delays contributed by peripheral processes, by response selection, and by stimulus identification combine additively in units of speed.

Since both stimulus surprisal and conditional response surprisal are measured in the same units (in bits), and since the slope coefficients for these two variables were similar in our data (see Table 5), these coefficients may turn out to be identical. Such a possibility is important,

Table 4. Across-words correlations of conditional response uncertainty and stimulus surprisal (or number of reproductions) with mean speed; also, conditional response uncertainty with stimulus surprisal (or number of reproductions)

	Uncertainty with speed	Surprisal (repro- ductions) with speed	Uncertainty with surprisal (repro- ductions)
English-speaking samples			
Male employees ($N = 12$)	-.69	-.37	.10
Female employees ($N = 14$)	-.61	-.38	.09
Male nonpsychotics ($N = 10$)	-.62	-.41	.10
Female nonpsychotics ($N = 22$)	-.72	-.38	.09
Male psychotics ($N = 9$)	-.48	-.37	.10
Female psychotics ($N = 4$)	-.49	-.35	.09
Norwegian-speaking samples			
Normal recruits ($N = 35$)	-.78	(.61)	(-.52)
Rejected recruits ($N = 35$)	-.73	(.53)	(-.50)

Table 5. Summary of the regression of mean response speed on conditional response uncertainty and stimulus surprisal; English-speaking samples only

	Multiple correlation	Intercept	Unstandardized regression weights	
			Uncertainty	Surprisal
Male employees	.74	1.14	-.09	-.06
Female employees	.69	1.25	-.08	-.07
Male nonpsychotics	.71	1.15	-.08	-.07
Female nonpsychotics	.79	.96	-.07	-.05
Male psychotics	.58	1.06	-.07	-.07
Female psychotics	.58	.79	-.05	-.05

Note: Sample N s as in Table 4.

since it would allow a major simplification of Equations 4 and 5, and imply that the effects of stimulus unexpectedness and overt response competition are directly additive in bits without prior differential weighing. Although the conclusion that one bit of uncertainty has the same magnitude of effect regardless of whether it is associated with stimulus-unexpectedness or with response-selection effects would be surprising, it cannot be ruled out here.

Stimulus surprisal and conditional response uncertainty

The intraindividual correlation between conditional response surprisal and stimulus surprisal was positive for 51 of 72 English-speaking

subjects. Although this permits rejection of the zero-correlation hypothesis (binomial test, $p < .05$), the means of the intraindividual correlations (Table 3) and the correlation between stimulus surprisal and conditional response uncertainty (top part of Table 4) were low. This lends weak support to Stilson's (1966) suggestion that unexpected stimuli produce more variable activity in the central nervous system and, consequently, less predictable response selection.

Conditional response uncertainty and number of reproductions

The Norwegian-speaking subjects were asked to reproduce their word-association responses. A simple model of this process assumes that response selection is simply a random choice according to the probabilities $P(R_i|S_j)$ and that the first and second response choices are independent. Then the probability that some response to stimulus S_j will be reproduced is simply

$$\sum_{i=1}^{r_j} P(R_i|S_j)^2.$$

Now, it can be shown that if $H(R|S_j) > H(R|S_k)$, then $\sum P(R_i|S_j)^2 < \sum P(R_i|S_k)^2$. That is, the probability of response reproduction decreases as conditional response uncertainty increases.

The correlations between the number of subjects who reproduced their response to each stimulus word and estimates of conditional response uncertainty were $-.50$ and $-.52$ for the two Norwegian-speaking samples (Table 4), figures comparable to Laffal's (1955) finding of a correlation of $.59$ between conditional response uncertainty and number of reproduction failures. Although the model suggested above seems exceedingly simple, these correlations are consistent with it. Also, as one might expect from the preceding results, mean response speed and number of subjects correctly reproducing their responses were positively correlated in the two Norwegian-speaking samples (Table 4).

It has been proposed (Stilson, 1966) that a single neurophysiological mechanism accounts for both the slowness and the looseness of association (in the sense of improbable associations) of schizophrenics. And if one adopts Cameron's (1947) continuity principle, the assumption that psychopathology varies continuously from excellent mental health to severe psychotic disorganization, then it follows that individual differences in slowness and looseness of association should covary. The within-groups correlations between mean response speed

and mean conditional response surprisal were $-.46$ in the English-speaking samples and $-.36$ in the Norwegian-speaking groups ($p < .05$ in both instances, see Table 6).⁴

Table 6. Summary of the across-persons correlations of mean speed with mean conditional response surprisal, mean speed with number of reproduction failures, and mean response surprisal with number of reproduction failures

	Speed with surprisal	Speed with reproduction failures	Surprisal with reproduction failures
English-speaking samples			
Male employees	$-.22$		
Female employees	$-.46$		
Male nonpsychotics	$-.42$		
Female nonpsychotics	$-.72^*$		
Male psychotics	$-.19$		
Female psychotics	$-.25$		
Norwegian-speaking samples			
Normal recruits	$-.34$	$-.16$	$.36^*$
Rejected recruits	$-.38^*$	$-.48^{*a}$	$.04^a$

Note: Sample N s as in Table 4, except as shown at ^a, where the N s were 19.
* $p < .05$.

Stilson (1966) has also suggested that a single defect accounts for the slowness and looseness of association among schizophrenics and explains their inability to reproduce their word-association responses. Invoking the continuity principle again leads to the expectation that mean speed and number of reproductions are positively correlated. The mean within-groups correlation between these variables was $.27$ (Table 6), which provides weak support for this expectation.

According to the response-reproduction model outlined above, an improbable response is less likely to be given twice than is a more probable one. This leads to the prediction of negative covariation between mean conditional response surprisal and number of reproductions. The observed across-persons correlation between the means of these two variables was $-.25$ (Table 6). This provides weak support for the hypothesis.

DISCUSSION

Relationship to previous studies

Our results are consistent with those of previous studies showing

a monotonic decrease in response latency with increasing response frequency (Wreschner, 1907; Menzerath, 1908; Cason and Cason, 1925; Schlosberg and Heineman, 1950). They are also compatible with the two studies that reported latency to be an increasing function of response diversity (Crane, 1915; Laffal, 1955). In only one case, however, is a more detailed comparison possible. Since Thumb and Marbe (1901) published the individual responses of their subjects, we were able to calculate the correlation between mean latency and an estimate of conditional response surprisal. This correlation of .90 is smaller in absolute value than that of $-.96$ between mean *speed* and conditional response surprisal for their data. Thus, our results for English and Norwegian are similar in their implications to those obtained by Thumb and Marbe for German at the turn of the century. Taken as a whole, this suggests that the relationship of speed and conditional response uncertainty is a robust one. And finally, our finding of a linear relationship between response speed and stimulus surprisal is consistent with the results of Hall and Ungelow (1957).

Generalized Marbe's law

Equations 4 and 5 can be viewed as a more general and more precise form of Marbe's law. Whereas the original statement was that response latency decreased monotonically with increasing response frequency (Thumb and Marbe, 1901), the present formulation incorporates the effect of stimulus unexpectedness as well, and thereby increases the generality of Marbe's law. Moreover, the present formulation suggests that the response-frequency effect may be the result of response competition, a possibility that becomes more apparent when response-frequency effects are measured by conditional response uncertainty.

In the form suggested here, Marbe's law becomes more precise than it was originally. Rather than the relatively weak statement of a monotonic relationship between latency and frequency, the present formulation specifies that certain transformations of these variables are *linearly* related. This, of course, greatly increases the empirical vulnerability of the proposition and narrows the variety of plausible models of the word-association process.

Model of the word-association process

At this point, we sketch a model of what may serve as the mechanism of word-association behavior. To begin with, let us distinguish two

aspects of the probabilistic structure of the task. The first of these, stimulus-word surprisal, is thought to reflect stimulus unexpectedness and to affect the speed of the recognition process. The second aspect concerns conditional response surprisal, its conditional mean $H(R|S_i)$, or its mean $H(R|S)$, all of which are considered measures of response-competition effects that influence the speed of the response-selection process.

One interpretation of stimulus unexpectedness assumes that the less probable the stimulus word, the greater the strength of the orienting reaction (Berlyne, 1960; Sokolov, 1960). If the orienting reaction's primary effect on the central nervous system is indeed to disrupt ongoing activity in preparation for processing new information, as suggested by Hunt (1963), then the less likely the stimulus, the greater the disruption (or the longer the recovery time) and the longer the response latency. Models based on this sort of formulation have been described by Stilson (1966), and they appear to have some promise in accounting for variations in recognition latency as a function of stimulus surprisal.

Suppose the stimulus sets in motion a sequence of events in the central nervous system that is a composite of stimulus effects and intrinsic neural noise. These neural events lead to parallel accumulation processes, one associated with each response. Accumulation continues until the amount associated with one response exceeds that associated with every other one by at least some minimum amount. Then, efferent activity leading to that response begins.

The accumulation time that elapses before this activity begins is the central component of a response latency, and it will depend on several things. First among these is that the larger the number of competing responses (responses for which there is a nonzero accumulation rate), the longer it will take for the minimum-difference threshold to be exceeded. Second, the minimum-difference threshold for response initiation might be lower for more probable responses than for less probable ones. This means that more probable responses would be expected to occur with shorter latency. In general terms, at least, this provides a basis for the decrease in response speed and reproduction probability with increasing surprisal or conditional response uncertainty.

Perhaps such a formulation can provide a basis for understanding the prolonged word-association latency and response-reproduction failure that have been attributed to 'complex indicators' or, more generally, to 'emotionally loaded' words (see Laffal, 1955). It may be that

these words produce more neural noise in the sequence of central-nervous-system events than do neutral words and that the diffuseness of the central event sequence is therefore greater. Such an effect, which is similar to Hebb's (1949) mechanism for emotional disturbance, would increase the number of responses significantly affected by the stimulus word and decrease the magnitude of the effect on the most likely responses. And as we have seen, such increased response-competition effects are associated with longer latencies and reduced reproduction probabilities. Of course, such a mechanism should be invoked only after allowing for the response-competition effects reported above, effects which are presumably relatively free of dependence on emotional impact.

Notes

This work was begun while Stilson held a Special Research Fellowship (MH5416) and completed while he held a Career Scientist Award (MH5416) from the National Institute of Mental Health. It was supported in part by Public Health Service Research Grant MH11005 to the same author, by National Institute of Health Research Grant 00404-01 from the Division of Research Facilities and Resources, and by a grant from the Norwegian Research Council to Haaseth. Thanks are due to Gary Ball for assistance in the analyses of the data, to Frank Brieser for statistical advice, and to Joseph Ulehla, Philip Archer, and Strother Walker for helpful criticisms and suggestions. Author Walsmith is now at the Boeing Corporation, Seattle, Washington; Schneider is now at the Metropolitan State College, Denver, Colorado; and Rogers is now at Case Western Reserve University, Cleveland, Ohio. Revision received for publication August 27, 1971.

1. If f_{ij} is the frequency of occurrence of response i as a response to stimulus word j , N is the total number of responses to word j , and $\hat{r}_j$ is the number of distinct responses to word j , then the estimate of conditional response uncertainty for stimulus word j is

$$\hat{H}(R|S_j) = -\sum_{i=1}^{\hat{r}_j} \frac{f_{ij} + .5}{N} \log_2 \frac{f_{ij} + .5}{N},$$

since f_{ij} was 0 for some responses.

2. Using the notation of note 1, the conditional response surprisal of word i in response to stimulus word j was estimated by $-\log_2 [(f_{ij} + .5)/N]$.

3. There was a tendency for the bivariate distribution of some subjects to 'sag' in the middle when response speed was plotted as the ordinate. In considering this, it must be remembered that surprisal was estimated from response frequencies for a group of 250 subjects in the case of the English-speaking sample and 300 subjects in the Norwegian-speaking sample. Such a group undoubtedly gives a greater number of distinct responses than a single person has available in his repertoire, so our estimates of conditional response probability for an individual are probably low. This means that

conditional response surprisal is overestimated by an amount that increases with surprisal. It is possible that this accounts for the sag.

4. For most of the English-speaking subjects, data on four-choice reaction time were also available. The across-subjects correlation (within groups) between mean four-choice reaction time and word-association response speed was $-.47$ ($p < .05$). The correlation between mean four-choice reaction time and mean conditional word-association response surprisal was $.35$ ($p < .05$). Thus, it is not simply word-association speed that is related to looseness of association; rather, one suspects that general slowness and associative looseness covary, as Stilson (1966) has suggested.

References

- Atneave, F. 1959. *Applications of information theory to psychology*. New York: Holt.
- Baker, L. M., and Elliott, D. N. 1948. Controlled and free association-times with identical stimulus- and response-words. *American Journal of Psychology* 61:535-539.
- Berlyne, D. E. 1957. Uncertainty and conflict: A point of contact between information theory and behavior-theory concepts. *Psychological Review* 64:329-339.
- Berlyne, D. E. 1960. *Conflict, arousal, and curiosity*. New York: McGraw-Hill. Pp. 34-37.
- Cameron, N. 1947. *The psychology of the behavior disorders*. New York: Houghton-Mifflin.
- Cason, H., and Cason, E. B. 1925. Associational tendencies and learning ability. *Journal of Experimental Psychology* 8:167-189.
- Crane, H. W. 1915. A study in association reaction and reaction time. *Psychological Monograph* 18(whole no. 80).
- Fraisse, P. 1963. La perception des mots: Etude sur les relations entre le seuil de reconnaissance et le temps de dénomination. In *Problemes de psycho-linguistique*, ed. J. Ajuriaguerra et al., pp. 169-178. Paris: Presses Universitaires de France.
- Fraisse, P. 1964. Relations entre le seuil de reconnaissance perceptive et le temps de réaction verbale. *Psychologie Française* 9:77-85.
- Haaseth, K. J. 1968. *Norske ord-assosiasjonstestnormer*. Oslo: Universitetsforlaget.
- Hall, J. F., and Ungelow, A. 1957. Free association time as a function of word frequency. *Canadian Journal of Psychology* 11:29-32.
- Hebb, D. O. 1949. *The organization of behavior*. New York: Wiley.
- Hunt, J. McV. 1963. Motivation inherent in information processing and action. In *Motivation and social interaction: Cognitive determinants*, ed. O. J. Harvey, pp. 35-94. New York: Ronald Press.
- Kent, G. H., and Rosanoff, A. J. 1910. A study of association in insanity. *American Journal of Insanity* 67:37-126.
- Laffal, J. 1955. Response faults in word association as a function of response entropy. *Journal of Abnormal and Social Psychology* 50:265-270.
- Menzerath, P. 1908. Die Bedeutung der sprachlichen Geläufigkeit ider der formalen sprachlichen Beziehung für die Reproduktion. *Zeitschrift für Psychologie* 48:1-95.

- Oldfield, R. C., and Wingfield, A. 1965. Response latencies in naming objects. *Quarterly Journal of Experimental Psychology* 17:273-281.
- Palermo, D. S., and Jenkins, J. J. 1964. *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press.
- Schlosberg, H., and Heineman, C. 1950. The relationship between two measures of response strength. *Journal of Experimental Psychology* 40:235-247.
- Sokolov, Y. N. 1960. *Perception and the conditioned reflex*. New York: Pergamon.
- Stilson, D. W. 1966. Models of some elementary behaviors distinguishing schizophrenics from others. Paper presented to the Fourth World Congress of Psychiatry, Madrid.
- Thorndike, E. L., and Lorge, I. 1944. *The teacher's word book of 30,000 words*. New York: Columbia University Teachers College.
- Thumb, A., and Marbe, K. 1901. *Experimentelle Untersuchungen über die psychologischen Grundlagen der Sprachlichen Analogiebildung*. Leipzig: Wilhelm Engelmann.
- Wreschner, A. 1907. Die Reproduktion and Assoziation von Vorstellungen. *Zeitschrift für Psychologie*, Ergänzungsband 3:1-599.

continued from page 548

- Lachman, Sheldon J. *Psychosomatic Disorders: A Behavioristic Interpretation*. New York: Wiley, 1972. Pp. xi, 208.
- Lana, Robert E., and Rosnow, Ralph L. *Introduction to Contemporary Psychology*. New York: Holt, Rinehart and Winston, 1972. Pp. ix, 486.
- Lana, Robert E., and Rosnow, Ralph L. *Readings in Contemporary Psychology*. New York: Holt, Rinehart and Winston, 1972. Pp. viii, 341.
- Landauer, T. K. *Psychology: A Brief Overview*. New York: McGraw-Hill, 1972. Pp. 360. \$8.95.
- Laszlo, Ervin. *The Systems Views of the World*. New York: George Braziller, 1972. Pp. viii, 131. \$5.95; paperback, \$2.45.
- Lazarus, Arnold A., ed. *Clinical Behavior Therapy*. New York: Brunner/Mazel, 1972. Pp. vii, 230. \$8.50.
- Levi, Lennart, ed., *Society, Stress, and Disease*. Vol. 1, *The Psychological Environment and Psychosomatic Disease*. New York: Oxford University Press, 1972. Pp. xvi, 485. \$29.95.
- Levy, Ronald B. *Self-revelation through Relationships*. Englewood Cliffs, N. J.: Prentice-Hall, 1972. Pp. xii, 150. \$7.95; paperback, \$3.95.
- Lewis, Howard R., and Lewis, Martha E. *Psychosomatics: How Your Emotions Can Damage Your Health*. New York: Viking Press, 1972. Pp. xv, 335. \$8.50.
- Lewis, William C. *Why People Change; The Psychology of Influence*. New York: Holt, Rinehart and Winston, 1972. Pp. xii, 241.
- Lippman, Leopold. *Attitudes toward the Handicapped: A Comparison between Europe and the United States*. Springfield, Ill.: Charles C Thomas, 1972. Pp. xv, 118. \$9.25.
- Lubar, Joel F. *A First Reader in Physiological Psychology*. New York: Harper & Row, 1972. Pp. viii, 344. Paperback, \$4.95.
- MacInnis, Donald E. *Religious Policy and Practice in Communist China*. New York: Macmillan, 1972. Pp. xxiv, 392. \$7.95; paperback, \$3.95.
- McCall, Robert B., and Havemann, Ernest. *Study Guide with Programmed Units*. 2d ed. Chicago: Harcourt Brace Jovanovich, 1972. Pp. vi, 185.
- McGaugh, James L., ed. *Behavioral Biology*. New York: Academic Press, 1972. Pp. 142.
- McMahon Frank B., ed. *Psychology: Perspectives on the Hybrid Science*. Englewood Cliffs, N.J.: Prentice-Hall, 1972. Pp. ix, 285. Paperback, \$3.95.
- McMahon, Frank B. *Psychology, the Hybrid Science*. Englewood Cliffs, N.J.: Prentice-Hall, 1972. Pp. x, 629. \$9.95.
- Markel, Lester. *What You Don't Know Can Hurt You*. Washington, D.C.: Public Affairs Press, 1972. Pp. vi, 288. \$7.50.
- Martin, William T. *Writing Psychological Reports*. Springfield, Ill.: Charles C Thomas, 1972. Pp. 184. \$10.50.
- Mason, Edwin. *Collaborative Learning*. New York: Agathon Press, 1972. Pp. 215. \$6.95.
- Mathiesen, Thomas. *The Defences of the Wark: A Sociological Study of a Norwegian Correctional Institution*. New York: Harper & Row, 1972. Pp. ix, 259. \$5.50; paperback, \$2.50.

to be continued in March issue

REPORTING OF FRAGMENTATIONS IN THE BINOCULAR RIVALRY OF CONTOURS

Rosemary Cogan and Alvin G. Goldstein

Texas Tech University and University of Missouri

Male and female college students reported changes for 5 min while observing rivalry or pseudo rivalry (sequences of rivalrylike stimulus changes) with or without fragmentations, under rapid-rate, natural-rate, or slow-rate instructions; first- and second-grade girls had only natural-rate instructions. Differences in rivalry rates were not related to response mechanisms and may be presumed to reflect perceptual differences. Age, sex, instruction, and temporal variables affected rivalry independently.

Stimulus fragmentations, or partial disappearances of one or both stimuli in binocular rivalry, are not limited to a single stimulus class used in rivalry but have been reported in response to colored stimuli (Fechner, 1861; Hering, 1964; Meenes, 1930; Washburn, Faison and Scott, 1934; Asher, 1954; Meredith and Meredith, 1964; Fox, 1968) and in response to a variety of stimuli of intersecting contours (Du-Tour, 1760 [sic]; Renshaw, 1940; Barany and Hallden, 1947; Crain, 1961; Fox, 1963; Meredith and Meredith, 1964; Cogan, 1966, 1970; Cogan and Goldstein, 1967; Smith, 1968; Whittle, Bloor, and Pocock, 1968). At least in rivalry of contours, stimulus fragmentation is not of minor importance, since fragmentations represent some 30% of all perceptual changes (Smith, 1968; Whittle, Bloor, and Pocock, 1968; Cogan, 1970) and all subjects report the appearance of fragmentations during contour rivalry when asked (Cogan, 1966).¹

When subjects are asked in typical rivalry studies to report the perception of either half-image or of both half-images together, the occurrence of fragmented perceptions may create a serious reporting problem. For instance, when viewing a vertical and horizontal line as half-images, a subject may perceive a sequence of perceptual changes such as

— ⊥ + ⊥ + .

He then has several equally reasonable reporting alternatives, ranging from 'horizontal, both' to 'horizontal, both, horizontal, both' or perhaps 'horizontal, vertical, both, horizontal, both.'

The question here is whether such difficulties in reporting fragmented perceptions might be related to several rather puzzling sets of findings within the rivalry literature. Instructions affect rivalry rates when subjects are asked to try to increase, decrease, or not alter the rate of rivalry changes (Southall, 1925; McDougall, 1903; Fry, 1936; Renshaw, 1940; Miller, 1958; Meredith and Meredith, 1962, 1963; Lack, 1967, 1969). Children report fewer rivalry changes than young adults (Murroughs and Ball, 1951; Goldstein and Cofoid, 1965). Men report fewer rivalry changes than women (Goldstein and Cofoid, 1965). Finally, rivalry rates increase during the first several minutes of viewing interrupted by rest periods (Gramlich, 1948; Goldstein and Cofoid, 1965; Aafjes, Hueting and Visser, 1966; Cogan and Goldstein, 1967), while a smaller increase in rivalry rate is typically found during continuous or uninterrupted viewing (Aafjes, Hueting, and Visser, 1966; Cogan and Goldstein, 1967).

It seems possible that in each of the above instances, differences in strategies adopted in *reporting* ambiguous fragmented perceptions might account for the differences found. Perhaps instructions act simply to establish the criterion to be adopted in reporting fragmented perceptions. Perhaps children tend to interpret their instructions more literally than adults and tend not to report fragmented perceptions. Similarly, sex and temporal effects on rivalry might reflect differences in reporting the rivalry experience rather than differences in the actual *perception* of rivalry.

Although rivalry perceptions are private experiences not directly accessible to public inspection, evaluation of the 'reporting' hypothesis is possible. If subjects were presented with visual stimuli that simulated real rivalry along various parameters (a pseudorivalry situation), within which the presence of fragmentation sequences might be manipulated, responses to real rivalry could then be compared with responses to pseudorivalry situations that either included or did not include fragmentation sequences. It would be possible to infer, in this manner, the way in which fragmentation sequences known to occur in contour rivalry are handled by the subjects.

If the findings discussed earlier result from different solutions to the problem of reporting fragmented perceptions, some rather clear results may be expected. If the 'reporting' hypothesis is correct, the responses of subjects in a pseudorivalry viewing condition that includes fragmentations should be like the responses of subjects viewing real rivalry; while the reports of subjects in a pseudorivalry condition without fragmentations should differ from reports of subjects viewing

real rivalry and should not be affected by variations in age, sex, instructions, or viewing time. If the 'reporting' hypothesis is not correct, the responses of the subjects in *both* pseudorivalry conditions should differ from the reports of subjects viewing real rivalry but should not be differentially affected by variations in age, sex, instructions, or viewing time.

Preliminary work (Cogan, 1970) indicated that when attention is paid to various perceptual units in rivalry, the frequency and duration of perceptions are affected. The problem of developing pseudorivalry sequences resembling real rivalry was approached by using experienced subjects able to continuously report their perception of rivalry changes.

METHOD

Development of pseudorivalry sequences

—Subjects—Two men and one woman served as subjects. Each had been in an earlier rivalry experiment, and each had normal vision as evaluated by Bausch & Lomb Ortho-Rater measures of acuity and phoria at near and far viewing positions.

—Stimuli—The stimuli consisted of a monocular vertical and a monocular horizontal line. Each line was 1.6 cm long by .1 cm wide and subtended 3.9 deg of visual angle. Each was surrounded by two fusion circles measuring 1.8 cm and 2.2 cm in diameter respectively, with a .1-cm line width. Each monocular stimulus field subtended 5.6 deg of visual angle.

—Apparatus—The stimuli were projected as white figures on a black field with two Grason Stadler E4580 projectors mounted on a Keystone telebinocular modified so that interstimulus spacing could be adjusted. Viewing distance was 23 cm. The extraneous visual field was masked with flat-back construction paper mounted around the projectors, exposing the circular stimulus fields to the subject's view.

—Procedure—Each subject refamiliarized himself with the appearance of rivalry and practiced reporting the following four categories of rivalry simultaneously:

- | | | |
|------------|--|---|
| Category 1 | Both images complete | + |
| Category 2 | One image complete, with part of the second image, e.g., | + |
| Category 3 | One complete image alone, e.g., | l |
| Category 4 | Parts(s) of one or both images, e.g., | + |

Each subject responded by using one of four response keys next to the telebinocular to report one of the four categories of perceptions. Subjects reported rivalry perceptions continuously with no evident difficulties for five 1-min viewing periods with 30-sec rest periods between each viewing minute.

—Construction of sequences—The pseudorivalry sequence *with fragmentations* was constructed by sampling five 1-min viewing periods at random from the 15 viewing minutes available. In all cases, fragmentations were arbitrarily limited to half-lines for practical reasons. Within category 2 each of the many possible combinations was assumed to occur with equal probability. Since category 4 presents many possible orientations but occurs infrequently, it was represented by only two of its possible representations, randomly selected.

To construct the pseudorivalry sequence *without fragmentations* the same 5 min of reporting were used. Only categories 1 and 3 were to be represented here. Since these categories have been found (Cogan, 1970) to occur with equal probability, category 1 or 3 was randomly assigned to each of the durations in the minutes selected for the pseudorivalry sequence with fragmentations. Thus, the frequency and duration of stimulus sequences were identical in both pseudorivalry sequences, as can be seen in Table 1, although four categories were represented in the fragmentation conditions and two categories were represented in the no-fragmentation conditions.

Table 1. Sequences of stimulus changes in conditions of pseudo rivalry with fragmentations and of pseudo rivalry without fragmentations; first 1-min viewing period

Stimulus sequence		Duration (sec)	Stimulus sequence		Duration (sec)
With fragmentations	Without fragmentations		With fragmentations	Without fragmentations	
1. +	+	1.0	15. -		2.0
2. +	—	.3	16.	—	2.3
3. +	+	5.2	17. +		.7
4. +	—	.8	18. —	+	2.7
5. +		2.0	19. +		5.0
6. +	+	.5	20. -	—	2.5
7. +	—	.9	21. +	+	1.7
8. —		6.1	22.	—	1.2
9. +	+	1.6	23. +		1.2
10. +	—	.6	24.	—	2.8
11.	+	3.8	25. +	+	.8
12. +		1.5	26.	—	2.5
13. —	—	2.2	27. +	+	7.5
14. +	+	.6			

Evaluation of the 'reporting' hypothesis

—Subjects—One hundred and seven students from Texas Tech University volunteered as subjects. Sixteen students were discarded from the sample because of visual anomalies, and one student was discarded who reported no rivalry during 5 min of viewing. This left 45 men and 45 women in the adult sample. Eighteen first- and second-grade girls also participated as

subjects. Three girls were discarded from the sample because of visual anomalies.

—Stimuli—The stimuli consisted of the projected images described above. Three variations of the sequences of stimulus changes were presented: *real rivalry*, in which the monocular vertical and monocular horizontal stimuli were presented simultaneously without change for 1-min intervals; *pseudo rivalry with fragmentations*; and *pseudo rivalry without fragmentations*, in which the pseudorivalry sequences were presented binocularly.

—Apparatus and procedure—The apparatus was that described above, except that key depressions were recorded on the moving paper of an Esterline Angus event recorder. Stimulus sequences were controlled with BRS solid-state and relay equipment and a BRS paper-tape timer.

After the rivalry situation was described in general terms, each subject was asked to report his perceptions of one viewing condition by depressing one or both of two keys positioned by the telebinocular. Each was asked to depress the right key when a vertical line was seen, the left key when a horizontal line was seen, and both keys when both lines were seen. Each practiced by responding to whole horizontal, vertical, and both-line stimuli drawn onto 5-by-7-in. cards and presented by the experimenter at the side of the telebinocular until responses appeared stable to both the subject and the experimenter.

Each subject was asked to view and report rivalry for five 1-min periods of time, with 30-sec rest periods after each viewing minute. Each subject was then given one of three sets of instructions, adopted from those of Meredith and Meredith (1962). The *rapid-rate instructions* were: "While you are watching, I would like you to look at the figure in an active manner. Look steadily at the center and try to make the lines change as quickly as possible. You can do this by concentrating. If you see the vertical line, concentrate on seeing the horizontal line next; if you see the horizontal line, concentrate on the vertical line until you see it." The *natural-rate instructions* were: "While you are watching, I would like you to look at the figure in a passive manner. Look steadily at the center and do not try to control the changing of the lines. Just relax and let the lines change at their own 'natural' rate." The *slow-rate instructions* were: "While you are watching, I would like you to look at the figure in an active manner. Look steadily at the center and try to make the lines change as slowly as possible. You can do this by concentrating. If you see the vertical line, try to keep seeing it. If you see the horizontal line, concentrate on it and try to keep seeing it as long as you can."

Each adult was randomly assigned to one viewing condition and one set of instructions; viewing conditions were counterbalanced in order and instructional conditions assigned in random order within each viewing sequence. Each girl was randomly assigned to one viewing condition and all were given natural-rate instructions; viewing conditions were presented in counterbalanced order.

The number of rivalry changes and the number of rivalry cycles reported were tabulated for each subject during each 1-min viewing period.² In the two pseudorivalry sequences, where the rate of changes per minute varied across viewing time, the percentage of the changes and the percentage of

the cycles presented which were reported were tabulated for each subject in each 1-min viewing period. The data on number of changes, number of cycles, percentage of changes, and percentage of cycles were analyzed in two groups of repeated-measures analyses of variance: comparisons between adult data, and comparisons between the reports of girls and women.

RESULTS AND DISCUSSION

The two pseudorivalry sequences were used to evaluate the experimental hypothesis, which predicts (a) that similar changes in rivalry rate over viewing time should obtain for subjects viewing real rivalry or pseudo rivalry with fragmentations, while no changes in rate should occur over viewing time for subjects viewing pseudo rivalry without fragmentations; (b) that similar instructional effects should obtain for subjects viewing real rivalry and pseudo rivalry with fragmentations, while no instructional effects should obtain for subjects viewing pseudo rivalry without fragmentations; (c) that similar differences should be found between the rivalry rates of men and women among subjects viewing real rivalry and pseudo rivalry with fragmentations, while no differences should be found between men and women viewing pseudo rivalry without fragmentations; and (d) that similar differences between the rates of women and girls should be found among subjects viewing real rivalry and pseudo rivalry with fragmentations, while no differences should be found between the rates of women and girls viewing pseudo rivalry without fragmentations. In each instance, an interaction between the two pseudorivalry conditions is predicted by the experimental ('reporting') hypothesis.

Viewing conditions

More changes and cycles were reported with real rivalry than in either pseudorivalry situation [number of cycles, $F(2, 72) = 96.0, p < .001$; number of changes, $F(2, 72) = 21.3, p < .001$] and more changes and cycles were reported in the pseudorivalry situation without fragmentations than in the one with [percentage of changes, $F(1, 48) = 164.5, p < .001$; percentage of cycles, $F(1, 48) = 20.4, p < .001$]. Evidently, part of the difficulty presented by the appearance of fragmented stimuli that were ambiguous in terms of instructions was resolved by a tendency for subjects to report no change when fragmentations were presented.

Temporal variables

The present data replicate the previously reported increase in rivalry rate over several minutes of discontinuous viewing time (Aafjes et al., 1966; Cogan and Goldstein, 1967). In the present study, rivalry rate was found to increase over viewing time for all measures among adults [number of cycles, $F(4, 288) = 36.1, p < .001$; percentage of cycles, $F(1, 192) = 6.7, p < .001$; number of changes, $F(4, 288) = 66.3, p < .001$; and percentage of changes, $F(4, 192) = 18.8, p < .001$]; per-both age groups [number of changes, $F(4, 96) = 11.8, p < .001$; percentage of changes, $F(4, 64) = 6.5, p < .001$; number of cycles, $F(4, 96) = 6.6, p < .001$; and percentage of cycles, $F(4, 64) = 5.1, p < .001$]. The effect of viewing time on rivalry rate did not interact with other variables, variables which, as we shall see, also produced rivalry changes.

Inspection of Figures 1-4 indicates that rivalry rate tended to increase over viewing time for both real rivalry and pseudo rivalry without fragmentations, while the rate reported in response to pseudo rivalry with fragmentations remained rather stable over viewing time. This finding is reflected in the reliable interaction between viewing time and viewing conditions among adults [number of cycles, $F(8, 288) = 12.9, p < .001$; percentage of cycles, $F(4, 192) = 5.1, p < .001$; number of changes, $F(8, 288) = 14.5, p < .001$; and percentage of changes, $F(4, 192) = 13.5, p < .01$] and in the interaction between the two age groups, which approached reliability [number of changes, $F(8, 96) = 1.8, p < .09$; and percentage of changes, $F(4, 64) = 2.2, p < .08$].

The data just cited are in direct contradiction to the prediction that the reports to pseudo rivalry with fragmentation would resemble those to real rivalry. It seems clear that responses to fragmentation are not related to the time changes in rivalry rate that were under consideration here. In addition, the tendency of subjects viewing pseudo rivalry without fragmentations to report a greater percentage of the *whole-line* rivalry changes presented to this group over viewing time does not seem a likely explanation for response changes in rivalry rate over viewing time in real rivalry. Although subjects viewing pseudo rivalry without fragmentations did tend to report a greater percentage of whole-line rivalry changes over viewing time, the group seeing pseudo rivalry with fragmentations—a group also exposed to whole-line rivalry changes—showed no increase over viewing time. If the tendency to report more whole-line changes later in viewing time were critical, changes for *both* pseudorivalry groups should be like changes observed

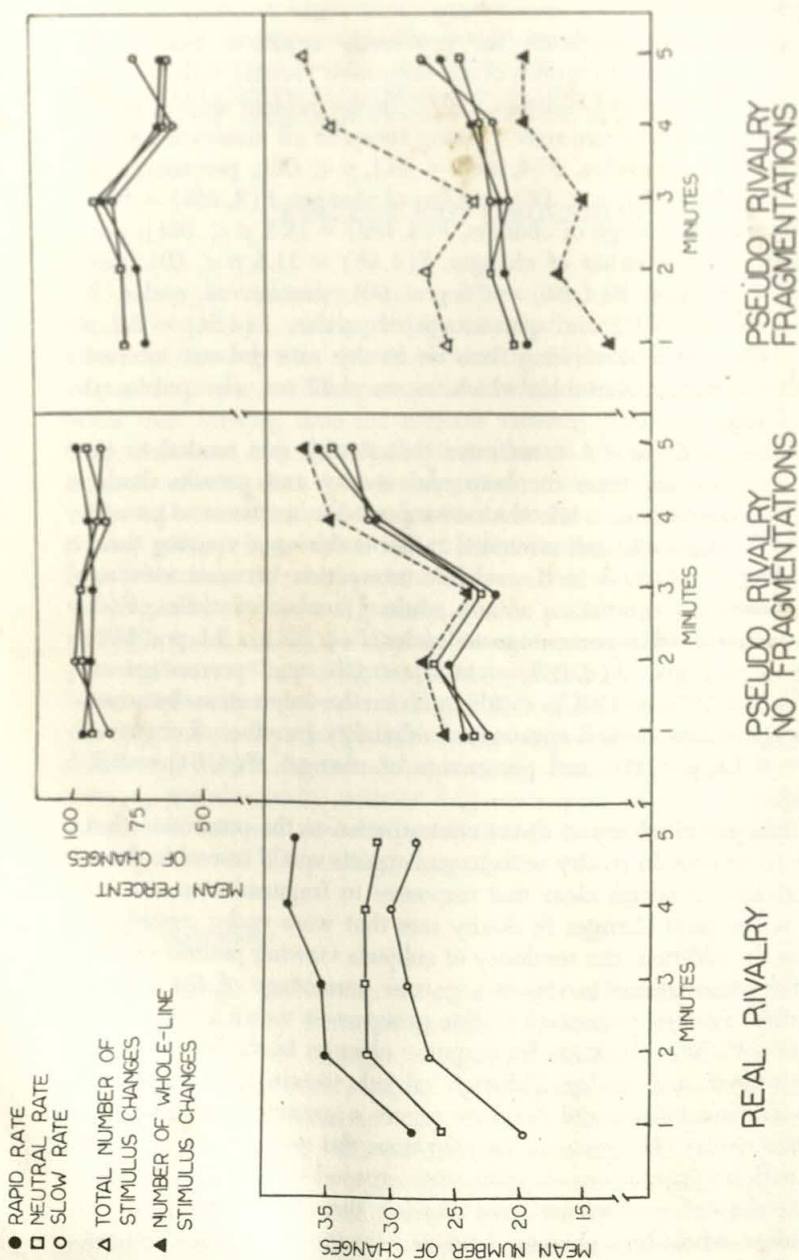


Fig. 1. Number and percentage of rivalry changes to real rivalry, pseudo rivalry without fragmentations, and pseudo rivalry with fragmentations as a function of instructions and viewing time

with real rivalry. The data suggest, then, that it is the perception of rivalry which changes over several minutes of spaced viewing in a real rivalry situation, and not simply the reports of what is perceived.

Instructional variables

The number of cycles reported was reliably affected by instructions among adults [$F(2, 72) = 3.25, p < .04$]. Interestingly, earlier literature has reported instructional effects for short viewing periods *only* when a 'cycles' measure of rivalry has been used (Miller, 1958). Only studies in which subjects have practiced the control of rivalry have reported instructional effects on dominance changes in rivalry (Meredith and Meredith, 1962, 1963; Lack, 1967, 1969), so there may be differences in the speed with which control of rivalry cycles and changes is achieved. The interaction between instructions and viewing conditions was reliable for the number of cycles [$F(4, 72) = 3.4, p < .01$] and approached reliability for the number of changes [$F(4, 72) = 2.00, p < .10$].

Although instructions affected rivalry rates among subjects viewing real rivalry (see Figures 1 and 2), no differences in rivalry rate were associated with instructional set for either pseudorivalry group. That is, the predicted interaction between instructional set and pseudorivalry viewing conditions was not found. It seems most probable that rivalry changes associated with instructional set do not arise from changes in the criteria for reporting perceptions of rivalry. The effect of instructions was similar for both men and women, was maintained throughout 5 min of viewing real rivalry, and was superimposed upon changes in rivalry rate associated with time.

Sex variable

The interaction between sex and viewing condition was reliable both for changes [$F(2, 72) = 3.16, p < .05$] and for cycles [$F(2, 72) = 3.44, p < .04$], and it can be seen in Figures 3 and 4 that the women viewing real rivalry reported more changes and cycles than did the men. The sex differences were also reliable for both percentage of changes [$F(1, 48) = 3.93, p < .05$] and percentage of cycles [$F(1, 48) = 15.13, p < .03$] within which the two pseudorivalry conditions were directly compared. Within the pseudorivalry conditions, as can be seen in Figures 3 and 4, men reported reliably more changes and cycles than women.

The reliable interaction between sex and pseudorivalry viewing

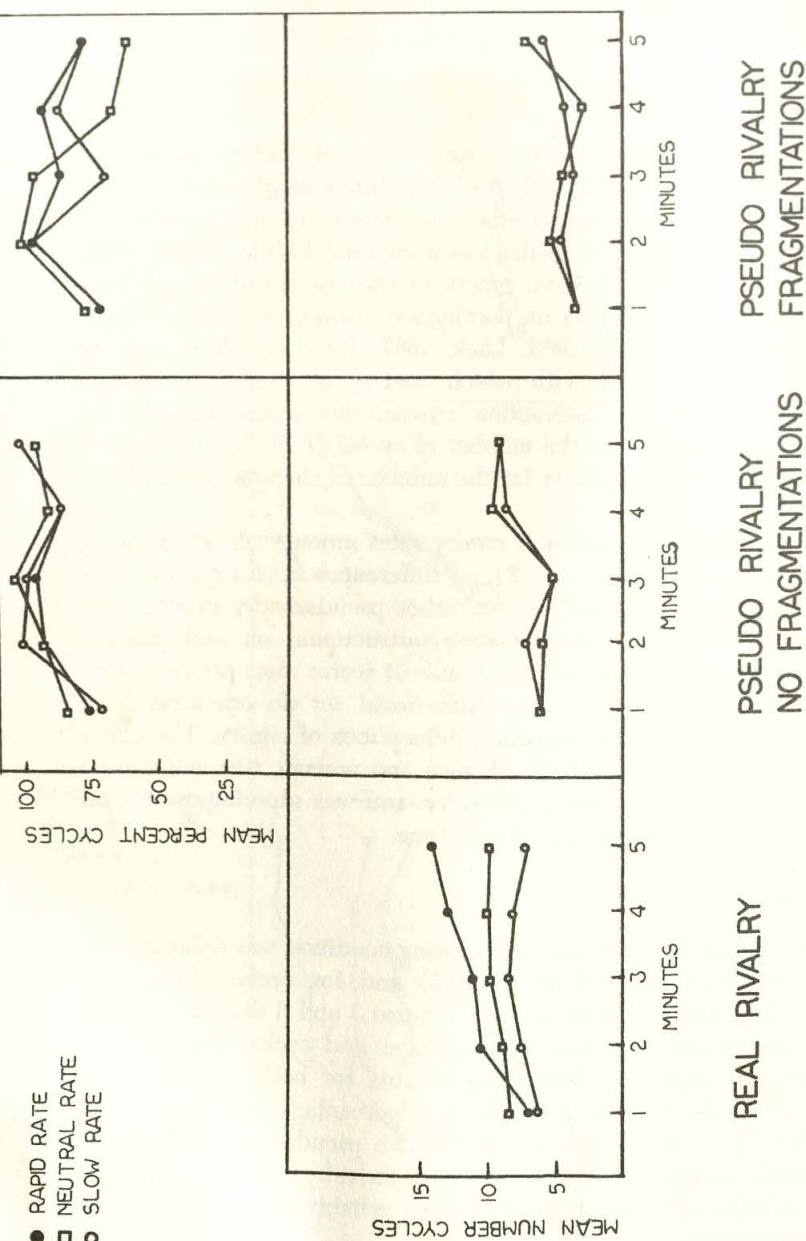


Fig. 2. Number and percentage of rivalry cycles to real rivalry, pseudo rivalry without fragmentations, and pseudo rivalry with fragmentations as a function of instructions and viewing time

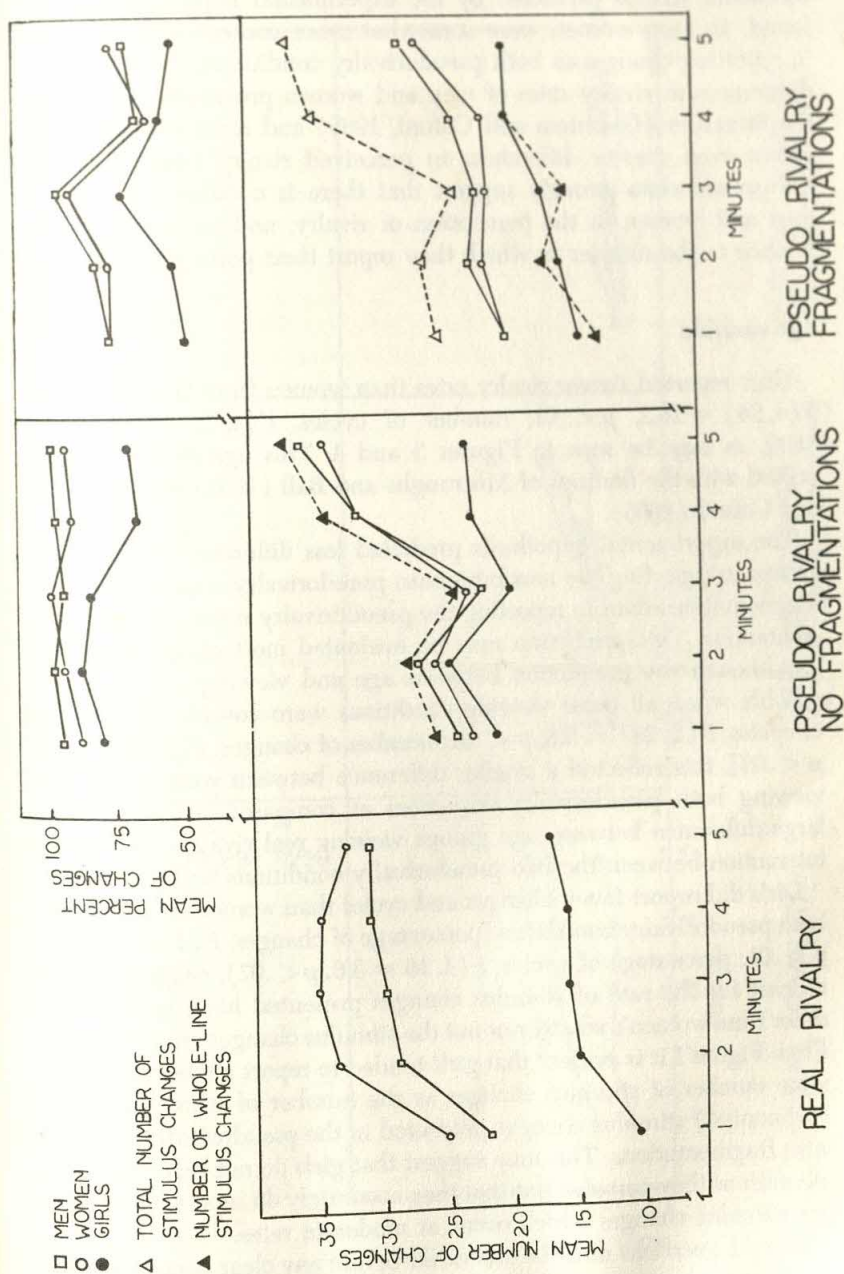


Fig. 3. Number and percentage of rivalry changes to real rivalry, pseudo rivalry without fragmentations, and pseudo rivalry with fragmentations as a function of sex, age, and viewing time.

conditions that is predicted by the experimental hypothesis was not found. In fact, women were somewhat more conservative than men in reporting changes in both pseudorivalry conditions, suggesting that differences in rivalry rates of men and women previously reported in the literature (Goldstein and Cofoid, 1965) and replicated here may reflect even greater differences in perceived rivalry rates. Certainly, the present data strongly suggest that there is a difference between men and women in the perception of rivalry, and not simply a difference in the manner in which they report their perceptions.

Age variable

Girls reported slower rivalry rates than women [number of changes, $F(4, 24) = 16.2$, $p < .01$; number of cycles, $F(4, 24) = 14.2$, $p < .001$], as may be seen in Figures 3 and 4. This age difference is in accord with the findings of Murroughs and Ball (1951) and Goldstein and Cofoid (1965).

The experimental hypothesis predicted less difference between age groups in reporting the nonambiguous pseudorivalry sequence without fragmentations than in reporting the pseudorivalry sequence with fragmentations. This prediction may be evaluated most clearly in Figure 3. Although the interaction between age and viewing condition was reliable when all three viewing conditions were considered [number of cycles, $F(2, 24) = 5.8$, $p < .01$; number of changes, $F(2, 24) = 4.6$, $p < .02$], this reflected a smaller difference between women and girls viewing both pseudorivalry sequences as compared with the rather large difference between age groups viewing real rivalry. The critical interaction between the two pseudorivalry conditions was not reliable.

Girls did report fewer changes and cycles than women when viewing both pseudorivalry conditions [percentage of changes, $F(4, 16) = 16.7$, $p < .01$; percentage of cycles, $F(4, 16) = 3.6$, $p < .07$], suggesting that, at least for the rate of stimulus changes presented here, girls' reports differ from women's whether or not the stimulus changes are ambiguous. From Figure 1 it is evident that girls tended to report almost exactly the same number of stimulus changes as the number of whole-line (nonfragmented) stimulus changes presented in the pseudorivalry sequence with fragmentations. This may suggest that girls do not report the fragmentations they perceive but that they accurately do report nonambiguous stimulus changes which occur at moderate rates.

Several questions must be answered before any clear conclusions can be drawn about the relationship between differences in reporting

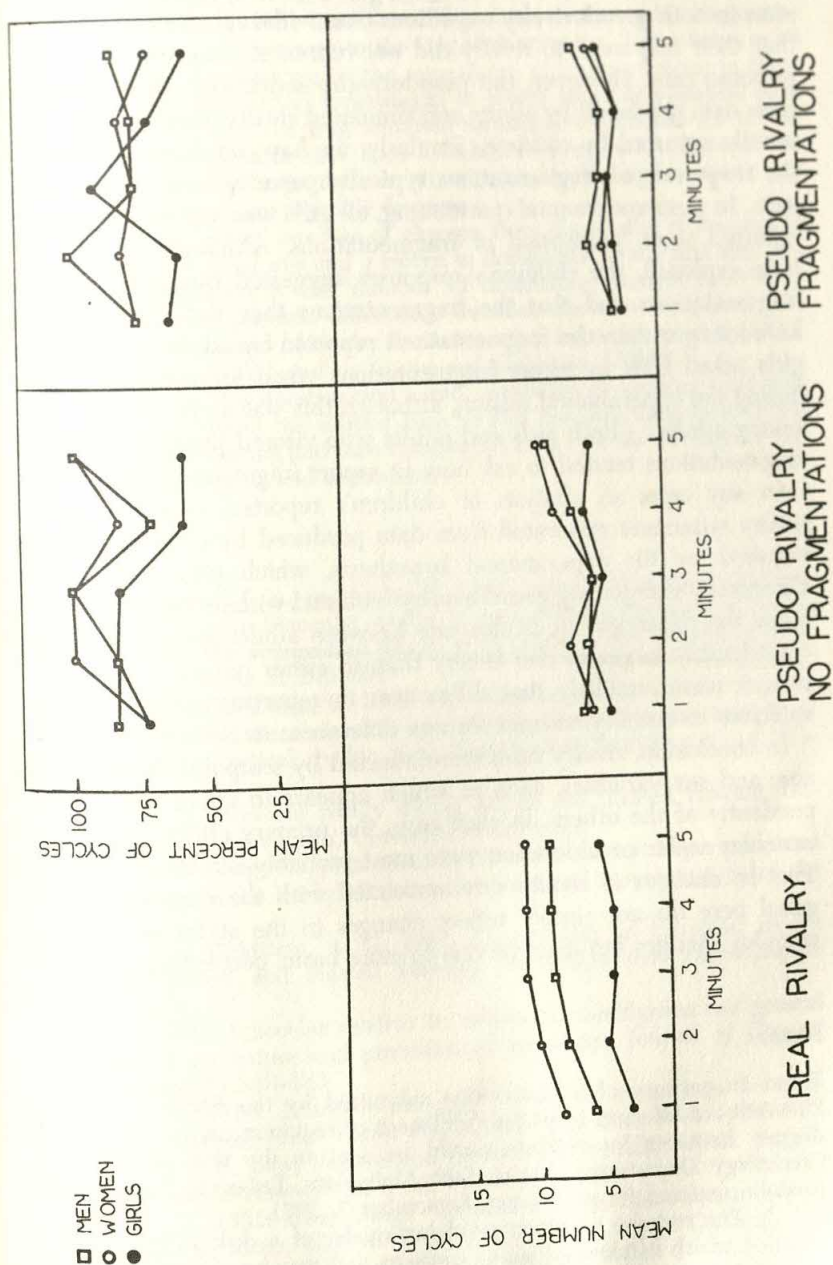


Fig. 4. Number and percentage of rivalry cycles to real rivalry, pseudo rivalry without fragmentations, and pseudo rivalry

perceptions of rivalry and the age variable. Children did report faster rates in both pseudorivalry conditions than with real rivalry, suggesting that their responses to rivalry did not represent a maximum possible response rate. However, the pseudorivalry sequences were generated from data produced by adults and simulated rivalry at rates faster than usually reported by children. Similarly, we have no direct estimate of the frequency of fragmentations typically perceived by younger subjects. In postexperimental questioning, all girls who viewed real rivalry reported some perception of fragmentations. Although this has not been explored, the children's responses suggested that they saw few fragmentations and that the fragmentations they did see may have lasted longer than the fragmentations reported by adults. None of the girls asked how to report fragmentations when viewing real rivalry during the experimental session, although this was a common question among adults.³ (Both girls and adults who viewed pseudorivalry with fragmentations tended to ask how to report fragmentations.)

In any case, an analysis of children's reported rates to pseudorivalry sequences generated from data produced by children was not required by the experimental hypothesis, which predicted *relative* differences in reporting pseudorivalry with and without fragmentations. Since the difference in rivalry rate between adults and children was considerably larger in real rivalry than in either pseudorivalry condition, it seems unlikely that differences in reporting styles would be sufficient to entirely account for age differences in rivalry rate.

In conclusion, rivalry rates were affected by temporal, instructional, age, and sex variables, each of which appears to affect rivalry independently of the others. In each case, the primary effects of the four variables under consideration were most probably directly perceptual. That is, changes in rivalry rate associated with the variables investigated here do not simply reflect changes in the styles of reporting stimulus changes but must be due to more basic, perceptual, processes.

Notes

Based on portions of a dissertation submitted by the first author to the University of Missouri in partial fulfillment of requirements for the doctoral degree. Requests for reprints should be sent to the first author at the Psychology Department, Texas Tech University, Lubbock, Texas 79409. Revision received for publication September 7, 1971.

1. The stimulus situation involving rivalry of a disk within a ring, a situation which involves adjoining contours and may be viewed as a special case of color rivalry, has provoked conflicting comments. Hering (1964)

observed alternation between a ring and a disk "whether in a limited area or throughout"; Levelt (1968) has taken the position that stimulus fragmentation does not occur in response to the ring/disk stimulus. There seem to be no experimental data bearing on the appearance of fragmentations in ring/disk rivalry, and in fact, no experiments in which stimulus variables have been manipulated and the fragmentation sequences evaluated. All of the variables under study here have been investigated previously *only* with the rivalry of moderately large contour stimuli, in which fragmentations are known to occur with considerable frequency.

2. Both the 'cycles' measure of changes from dominance of one eye to dominance of the other eye and a return to dominance of the first eye, and the simpler evaluation of the 'number' of dominance changes have been included to detect possible differences between the two measures. The problem of measurement of rivalry fluctuations has been discussed by Goldstein and Cofoed (1965) and by Cogan and Goldstein (1967).

3. Whenever this question was asked, regardless of viewing condition, the experimenter asked the subject how he had handled these perceptions during his earlier viewing and then commented only that the strategy he had adopted was reasonable and satisfactory.

References

- Aafjes, M., Hueting, J. E., and Visser, P. 1966. Individual and inter-individual differences in binocular rivalry in man. *Psychophysiology* 3:18-22.
- Asher, H. 1954. The suppression theory of binocular vision. *American Journal of Optometry* 31:246-251.
- Barany, E. H., and Hallden, U. 1947. The influence of some central nervous depressants on the reciprocal inhibition between the two retinæ as manifested in retinal rivalry. *Acta Physiologica (Scandinavia)* 14:296-316.
- Cogan, R. 1966. Binocular rivalry during spaced and massed viewing. Unpublished master's thesis, University of Missouri.
- Cogan, R. 1970. Image fragmentations in binocular rivalry. Paper presented at meetings of the Rocky Mountain Psychological Association, Salt Lake City, May.
- Cogan, R., and Goldstein, A. G. 1967. The stability of binocular rivalry during spaced and massed viewing. *Perception and Psychophysics* 2:171-174.
- Crain, K. J. 1961. Binocular rivalry: Its relation to intelligence and general theory of its nature and physiological correlates. *Journal of General Psychology* 74:259-283.
- DuTour, E. F. 1760. Pourquoi un objet sur lequel nous fixons les yeux paraît-il unique? *Mem. Math. Phys. Acad. Roy. Sci., Paris* 4:499-511. Cited by J. R. Law, Suppression theories of binocular vision, Unpublished doctoral dissertation, Duke University, 1967.
- Fechner, G. T. 1861. Über einige Verhältnisse des binocularen Sehens. *Abhandt. d. sachs. Ges. d. Wiss.* 7:337. Cited by S. Dawson, The experimental study of binocular colour mixture, *British Journal of Psychology*, 1916, 13:510-551.

- Fox, R., Jr. 1963. An analysis of the suppression mechanism in binocular rivalry. Unpublished doctoral dissertation, University of Cincinnati.
- Fox, R. 1968. Binocular rivalry and reciprocal inhibition. Paper presented at meetings of the Midwestern Psychological Association, Chicago, May.
- Fry, G. A. 1936. The relation of accommodation to the suppression of vision in one eye. *American Journal of Ophthalmology* 19:135-138.
- Goldstein, A. G., and Cofoid, D. 1965. A developmental study of retinal rivalry. *Perceptual and Motor Skills* 20:235-238.
- Gramlich, S. G. 1948. A study of the distribution of retinal rivalry rates and the relation of rivalry to some visual functions and skills. Unpublished master's thesis, Ohio State University.
- Hering, E. 1964. *Outlines of a theory of the light sense*, trans. L. M. Hurvich and D. Jameson. Cambridge, Mass.: Harvard University Press.
- Lack, L. 1967. Binocular rivalry control and selective attention. *Australian Psychologist* 2(1).
- Lack, L. C. 1969. The effect of practice on binocular rivalry control. *Perception and Psychophysics* 6:397-400.
- Levelt, W. J. M. 1965. Binocular brightness averaging and contour information. *British Journal of Psychology* 56:1-13.
- Levelt, W. J. M. 1968. *On binocular rivalry*. The Hague: Mouton.
- McDougall, W. 1903. The physiological factors of the attention process, III. *Mind* 12:473-488.
- Meenes, M. 1930. A phenomenological description of retinal rivalry. *American Journal of Psychology* 42:260-269.
- Meredith, G. M., and Meredith, C. G. W. 1962. Effect of instructional conditions on rate of binocular rivalry. *Perceptual and Motor Skills* 15:655-664.
- Meredith, G. M., and Meredith, C. G. W. 1964. Some attributive characteristics of binocular rivalry. *Perceptual and Motor Skills* 19:511-514.
- Miller, E. F. II. 1958 Evaluation of certain visual and related tests: V. Retinal rivalry. U.S.N. School of Aviation Medicine research report NM-140111, Sub 6, #5, II.
- Murroughs, T. R., and Ball, E. A. 1951. Relationship of retinal rivalry and stereopsis. *Optometric Weekly* 42:793-1795.
- Renshaw, S. 1940. Retinal or binocular rivalry. *Psychological Optics* 2(3).
- Smith, P. K. 1968. Image fragmentations with afterimages and in binocular rivalry. *Psychonomic Science* 10:275-276.
- Southall, J. P. C., ed. 1925. *Helmholtz's treatise on physiological optics*. Vol. 3, *The perceptions of vision*. Menasha, Wisc.: Optical Society of America.
- Washburn, M. F., Faison, C., and Scott, R. 1934. A comparison between the Miles ABC method and retinal rivalry as tests for ocular dominance. *American Journal of Psychology* 46:633-636.
- Whittle, P., Bloor, D. C., and Pocock, S. 1968. Some experiments on figural effects in binocular rivalry. *Perception and Psychophysics* 4:183-188.

RETENTION-TASK CARRYOVER EFFECTS AFTER VERBAL-DISCRIMINATION LEARNING

N. Jack Kanak, Lawrence E. Cole, and C. Dwayne Curtis

University of Oklahoma

Three experiments, each using a different recall procedure, failed to find reliable carryover effects between modified free-recall and associative-matching retention tasks either after verbal-discrimination learning of a single list (unlike Keppel, 1966) or in a nonspecific-transfer condition. The latter condition yielded significant unlearning of context-item associations in all three experiments and of incidentally learned specific associations in two.

Three experiments were conducted to clarify the methodological issue (raised by Keppel, 1966) of whether two within-subject measures of retention after verbal-discrimination learning result in serious confounding of the second measure. In addition, all three experiments investigated the unlearning of contextual and specific associations after practice on a nonspecific transfer list (Kanak and Curtis, 1970).

Keppel, in a study of single-list verbal-discrimination learning and retention, found that a counterbalanced, within-subject administration of two retention tasks—modified free *recall* (Briggs, 1954) and associative *matching* (McGovern, 1964)—resulted in significant carryover effects. When the matching task came first, scores on the recall task were inflated compared to when the recall task came first, a result which Keppel attributed to the opportunity for additional item learning during the prior matching task. More important to investigators who give first the recall, then the matching task in a within-subject design (as, for example, in retroactive-inhibition studies of the unlearning of response items and of specific associations) was Keppel's finding that scores on the matching task were depressed when it came second compared to when it came first. Keppel speculated that this effect was produced by the associative interference created by incorrect pairings on the prior recall task. Although the primary concern of each of the present experiments was with this effect of prior recall on matching-task scores, all three of the experiments employed a counterbalanced administration of both tasks, to test the replicability of Keppel's results when the recall task is given in a manner similar to or different from

that which Keppel himself used in his investigation in 1966.

The modified free-recall task is typically administered, whether in paired-associate or verbal-discrimination studies, in one of three ways. Separate subjects may be asked to recall *unidirectionally*, that is, to recall either in the forward or backward direction in response to presented items; *bidirectionally in discrete trials*, that is, to recall in a given direction and then in the opposite direction, with direction of first recall usually counterbalanced across subjects; or *unidirectionally in nondiscrete trials*, that is, to recall to a mixed list in which half of the cue items involve forward recall, the remainder backward recall.

Keppel's unpaced recall task involved the latter procedure, in a manner such that six wrong (W) and six right (R) items from the 12-pair verbal-discrimination list were presented for recall of the six R and W items, respectively, with which they had been paired. It is conceivable that Keppel's observed depression of matching-task scores after recall could be related to the mixed-list procedure and may not be a factor in other recall procedures. Kanak and Curtis (1970), however, found evidence which suggested that a similar effect might be operative with discrete-trial bidirectional recall. A significant negative correlation was found between errors or incorrect responses on the recall task and correct scores on a subsequent matching task. This relationship occurred only for a group which, as in Keppel's study, learned a single verbal-discrimination list that was followed immediately by the retention tasks; it did not occur either for a nonspecific-transfer group that learned an unrelated list 2 prior to retention tests on list 1 or for a 'normal-forgetting' control group. Kanak and Curtis concluded that the apparent lack of influence of prior recall on matching-task scores in the nonspecific-transfer group indicated that the matching-task measure of specific-association unlearning was, relative to the depression of scores in the immediate-recall group, a conservative estimate of retroactive inhibition.

Due to the importance of this methodological question to retroactive-inhibition research designs, Experiment I was designed to investigate the problem under conditions of unmixed-list unidirectional recall after verbal-discrimination learning. Experiment II tested the replicability of Keppel's findings, using (as he did) mixed-list, nondiscrete-trial unidirectional recall. Experiment III investigated discrete-trial bidirectional recall. In addition, each experiment included a nonspecific-transfer group to determine whether the potential carry-over effects extended to estimates of retroactive inhibition of context-item and specific associations.

EXPERIMENT I

Method

—Subjects—The 64 subjects were students in introductory psychology at the University of Oklahoma and participated as part of a course requirement. They were randomly assigned to the eight treatment groups (each $N = 8$) in the order of their appearance at the laboratory, within the restriction that a given block of eight cells (representing the eight treatment conditions) be filled before assignment of cell $n+1$. None of the subjects had previous experience in verbal-learning experiments.

—Materials—A pool of 60 familiar words was selected from the Palermo and Jenkins (1964) norms, with the primary criterion for selection being that all words be associatively unrelated to each other. Two subpools, each of 15 words, were randomly selected from the original pool, with one subpool randomly designated as R items and the other as W items. Random pairings were then generated to form list 1A. List 1B was formed by reversing the intrapair W and R functions of the items in list 1A. Lists 1C and 1D were constructed by randomly re-pairing lists 1A and 1B respectively, within the restriction that each W item now be paired with a different R item. The words for both treatment conditions—*single* list and *nonspecific-transfer* list—were the same as those used by Kanak and Curtis (1970), although that study did not involve counterbalancing of item function. The four variants of list 1 were employed to provide greater generality of the results and to minimize the influence of idiosyncratic associations, and each appeared equally often within each treatment condition.

List 2 for the *transfer* condition was formed from the remaining 30 words in the same manner as list 1A. Four random serial orders of each list were employed to minimize the influence of position cues. In addition, the left/right spatial position of R items was randomized within the restrictions that each R item appear equally often in a left/right spatial position and that runs of R items occupying a given spatial position be limited to four.

—Procedure—Before practice on the experimental list(s), each subject received standard instructions for verbal-discrimination learning, followed by a demonstration of the procedure (with materials unrelated to the experimental lists). The experimental lists were then presented on a Lafayette memory drum by the anticipation method at a 2:2-sec rate with a 4-sec intertrial interval. The paired items appeared in horizontal juxtaposition in both presentations, with the R item underlined during the feedback interval.

List 1 learning was carried to a criterion of one perfect trial; list 2 learning for the transfer group began immediately after attainment of the list 1 criterion and continued for ten trials. After the appropriate list, half ($N = 16$) of the subjects in the single-list and transfer conditions were given the recall task, then the matching task for list 1 pairs, while the other 16 subjects got the matching task first, then the recall task. In the recall task, half of the subjects ($N = 8$) in each of these four subgroups recalled R items in response to the W items printed on a sheet of paper, and the others recalled W items in response to R items. In the matching task, the W and R items

were presented in separate columns on a sheet of paper, and the subjects were instructed to match them. The stimulus materials for both retention tasks were presented in new random orders. As in Keppel's study, both retention tasks were unpaced and the subjects were instructed to guess if uncertain of the correct pairing. The between-group variables thus represent a $2 \times 2 \times 2$ factorial design with number of lists learned (single list versus transfer list), direction of recall (W-R versus R-W), and retention-task sequence (recall first versus matching first) as the three factors.

Results

In *acquisition*, the four variants of list 1 produced a range in mean trials to criterion of 7.00 to 8.94, differences which were clearly non-significant [$F(3, 48) = 1.54, p > .20$]. Furthermore, the main effects for the dummy variables of retention-task sequence [$F(1, 48) = 1.49$] and number of lists learned ($F < 1$) were also nonsignificant. The interaction of list variants and retention-task sequence, however, was significant [$F(3, 48) = 4.18, p < .025$]. No other interactions approached significance. The significant interaction resulted from the greater difficulty of lists 1A and 1C for subjects who later received the recall task first, while the 1B variant was somewhat more difficult for those who received the matching task first. These differences are most likely attributed to chance selection factors, since neither of the main effects approached significance.

That differential trials to criterion were associated with the subgroups in the interaction could represent a potential source of confounding and limit the generalizability of the results from the retention tasks. However, other studies (Kausler and Olson, 1969) have found negligible relationships between degree of learning and retention measures of incidental learning in the verbal-discrimination task. A similar conclusion seems warranted in the present study, since all of the eight product-moment correlation coefficients (range of $+.52$ to $-.46, df = 6$) between trials to criterion and recall scores for each combination of list variants and retention-task sequence, summed over direction of recall and number of lists learned, were nonsignificant. Only one of the coefficients resulted in a statistical trend ($p < .10$). In addition, all eight coefficients between trials to criterion and matching-task scores were nonsignificant (range of $+.39$ to $-.37, df = 6$). The conclusion that differential degrees of learning by the subgroups in the interaction did not measurably influence measures of retention therefore seems tenable. Consequently, the analyses reported below omit list variants as a factor.

Consistent with Keppel's 1966 report, the data on *modified free*

recall were treated by both stringent (only correct pairings counted) and lenient (total items recalled counted, without regard to correctness of pairing) scoring. Both types of analysis involved three factors: number of lists learned, retention-task sequence, and direction of recall, each at two levels.

Keppel reported that total item recall was significantly greater when the matching task came first, but it is not possible to determine from his report whether the same was true for only correct pairings, that is, for stringent as well as lenient scoring. In the present study, stringent scoring resulted in no effect of retention-task sequence ($\bar{X} = 6.16$ versus 6.17; $F < 1$), while lenient scoring produced a statistical trend toward better recall when the matching task was first than when the recall task was first ($\bar{X} = 8.46$ versus 6.97; $F(1, 56) = 2.94, p < .10$). In the latter analysis, however, retention-task sequence produced a trend toward a significant interaction with the number of lists learned [$F(1, 56) = 3.76, p < .10$]. No other interactions involving the retention-task sequence approached significance (all F s < 1).

The trend toward a significant interaction stemmed mainly from the transfer group, which recalled more items when the matching task came first (127 versus 78), while sequence did not affect the recall of the single-list group (145 versus 142). Since total item recall for the transfer group was significantly less than that for the single-list group [$\bar{X} = 6.46$ versus 8.97; $F(1, 56) = 9.35, p < .005$], the higher recall by the transfer group when the matching task came first likely reflects the reinstatement during the matching task of the availability of items that had undergone unlearning during list 2 as a result of competition between context-item associations of list 1 and list 2 (Kanak and Curtis, 1970).

The analysis of the stringent scores also indicated significantly poorer recall by the transfer group [$\bar{X} = 4.97$ versus 7.37; $F(1, 56) = 10.96, p < .005$], but the interaction with retention-task sequence was far short of significance [$F(1, 56) = 1.55, p > .10$]. It might also be noted that in both the stringent and lenient scoring analyses, W-R recall was significantly superior to R-W recall [$\bar{X} = 7.06$ versus 5.28 for stringent scoring, 8.66 versus 6.72 for lenient scoring; $F = 6.01$ and 5.35 respectively, both $ps < .025$], whereas direction of recall did not interact significantly with either of the other factors (all $ps > .25$).

In summary, then, although the analysis of total item recall produced a trend toward superior recall when the matching task was given as the first rather than as the second measure of retention, this superiority was limited to the transfer group and, in contrast to Keppel's findings,

did not extend to the single-list group. Furthermore, the sequence effect under stringent scoring did not even reach the level of a significant trend nor did it interact with the number of lists learned.

The analysis of scores for *associative matching* failed to produce any significant differences as a function of either sequence of retention tasks ($\bar{X} = 6.35$ for recall first versus 7.00 for matching first) or number of lists learned ($\bar{X} = 6.30$ for transfer versus 7.05 for control) and the interaction was nonsignificant (all F s < 1). In addition, the product-moment correlation coefficients between errors (both intra- and extra-list) on the prior recall task and correct scores on the subsequent matching task for the single-list and transfer groups, with subjects subdivided for each direction of prior recall, were all nonsignificant (critical $r_{.05} = \pm .62$, $df = 6$), although the value for the single-list, W-R recall group approached significance ($r = .61$; all other r s were between $-.40$ and $-.43$).

EXPERIMENT II

Experiment II employed a mix-list, nondiscrete-trial unidirectional recall task comparable to Keppel's (1966), and used the same materials as in Experiment I, in order to test the replicability of Keppel's findings with different subjects and a longer list, and to determine whether Keppel's observed depression of matching-task scores after mixed-list unidirectional recall is generalizable to the nonspecific-transfer condition.

Method

—Subjects—The 32 subjects were again students in introductory psychology at the University of Oklahoma and participated as part of a course requirement. None had been in any verbal-learning experiments, and they were assigned to the four treatment groups (each $N = 8$) as in Experiment I.

—Procedure—The verbal-discrimination lists and the conditions of presentation of each list were identical to those of Experiment I except that list 1 learning, in consonance with Keppel's procedure, was continued to two successive perfect trials. Ten trials were again given on the transfer list. Due to the uneven number of verbal-discrimination pairs (15), the mixed-list recall task was administered in such a way that half of the subjects in each of the four groups recalled eight W and seven R items to the eight and seven R and W items, while the remaining subjects recalled seven W and eight R items to the seven R and eight W items that were presented. The matching task was administered as in Experiment I. The major variables in Experiment II, then, were number of lists learned, retention-task sequence, and the within-subject factor of direction of recall.

Results

In *acquisition*, as was the case for Experiment I, the four variants of list 1 and the groups defined by number of lists learned and by retention-task sequence did not differ in mean trials to criterion ($F_s < 1$). In contrast to Experiment I, the interaction of list variants and retention-task sequence was nonsignificant ($F < 1$), and all other interactions fell far short of significance ($p_s > .20$). The four lists were therefore pooled within each of the four between-group conditions for all subsequent analyses.

For the data on *modified free recall*, to correct for the uneven number of instances of W-R and R-W recall, one W-R or R-W pair as appropriate was randomly eliminated from statistical consideration for all groups. The same pair, however, was eliminated in all groups. Unlike Keppel's results, the effect here for retention-task sequence was nonsignificant ($F_s < 1$) for both lenient scoring ($\bar{X} = 6.37$ for recall first versus 7.12 for matching first) and stringent scoring ($\bar{X} = 6.25$ for recall first versus 5.44 for matching first).

In fact, all main effects and interactions produced $F_s < 1$ except for the main effect of number of lists learned [$F(1, 28) = 4.60, p < .05$, for stringent scoring, versus 3.36, $p < .10$, for lenient scoring]. Unlearning of context-item associations for the transfer group relative to the single-list group was present, therefore, in both analyses, with respective means ($N = 16$) of 4.19 and 6.87 for stringent scores and of 5.87 and 7.62 for lenient scores. Also, contrary to the analyses of lenient scores in Experiment I, there was no statistical tendency supporting reinstatement of item availability during a prior matching task for the nonspecific-transfer group, although the absolute difference was in the expected direction. Unlike Experiment I, direction of recall was not significant under either type of scoring procedure.

Analyses of the data on *associative matching* for the 15 pairs indicated that the retention-task sequence again did not affect correct matching scores ($F < 1$), a result also inconsistent with Keppel's (1966). The means were 9.06 for recall first versus 8.37 for matching first. Product-moment correlations between errors in the recall task and correct scores in the subsequent matching task were not computed because of the extremely low incidence of incorrect pairings during recall, despite the instructions to guess when uncertain. The transfer group experienced significant unlearning of specific list 1 associations relative to the single-list group [$\bar{X} = 7.00$ versus 10.44, $F(1, 28) = 9.92, p < .005$]. The interaction of number of lists learned and retention-task sequence was nonsignificant [$F(1, 28) = 1.18, p > .25$].

EXPERIMENT III

Experiment III involved the discrete-trial bidirectional procedure used by Kanak and Curtis (1970) and the same materials as in Experiments I and II above. The overall experimental design involved a 4 (list variants) $\times$ 2 (number of lists learned) $\times$ 2 (retention-task sequence) $\times$ 2 (direction of recall) factorial, with the latter factor being a repeated measure.

Method

—Subjects—The 64 subjects came from the same source and were selected in the same manner as in Experiments I and II. None had been in any other verbal-learning experiments.

—Procedure—For half of the subjects in each cell, the recall task involved the recall of the 15 R items to the 15 W items listed on a sheet of paper, followed then by recall of W items to R items. For the other subjects, it was the counterbalanced, reverse sequence of recall. Otherwise, all procedural details were identical to those of Experiment II.

Results

For *acquisition*, the four list variants and the retention-task sequences both yielded $F < 1$ in the trials-to-criterion analysis. In addition, the single-list and transfer groups did not differ significantly [$F(1, 48) = 3.33, p > .10$]. However, the interaction of list variants and number of lists learned [$F(3, 48) = 3.33$] and the interaction of list variants and retention-task sequence [$F(3, 48) = 3.37$] were both significant ($ps < .05$). All other interactions yielded $F < 1$.

An examination of the product-moment correlation coefficients between trials to criterion and recall scores for the single-list and transfer groups, computed for the same reasons as in Experiment I, indicated only one significant correlation coefficient [$r(6) = +.72, p < .05$] among the eight cells of the interaction of variants and number of lists learned. This relationship involved one of the single-list cells. The correlations for the same cells between trials to criterion and matching-task scores resulted in two significant coefficients [$r(6) = -.70$ and $+.76, p < .05$], one each in the single-list transfer groups but both involving the same list variant as the one significant correlation in the recall data. The possible implications of these positive correlations involving this particular list variant must therefore be considered in the interpretations of the recall- and matching-task data, but the degree of learning in the remaining three list variants does not represent a poten-

tial source of confounding to the interpretation of the results in the single-list versus transfer group comparisons.

The same list variant produced one significant correlation ($r = +.66$) between trials to criterion and recall when the matching task came first, and two significant correlations ($r_s = +.70$ and $+.78$) on the matching data, one for each task sequence. Hence, these correlations must also be considered in interpreting the data on retention-task sequence, but the remaining seven and eight cells for the recall- and matching-task data, respectively, are not called into question with respect to the degree-of-learning problem.

Under stringent scoring of *modified free recall*, not only was the main effect for list variants nonsignificant ($F < 1$) but list variants did not interact with either number of lists learned or retention-task sequence ($F_s < 1$). In addition, none of the second-order interactions, or the third-order interaction, involving these factors approached significance. Comparable results were obtained for lenient scoring. Apparently, the positive correlations cited above in the trials-to-criterion analysis are not indicative therefore of any causal relationship between degree of learning and recall scores. The interpretation of the recall results for both stringent and lenient scoring with regard to the number of lists learned and retention-task sequence variables does not therefore appear to require qualification.

The main effect for retention-task sequence under lenient scoring yielded $F(1, 48) = 1.82$, $p > .10$, and $F < 1$ under stringent scoring. The means for lenient scoring were 16.25 for recall first and 18.22 for matching first, compared to 14.28 and 15.19 for stringent scoring. The second- and third-order interactions involving retention-task sequence for both lenient and stringent scoring typically yielded $F_s < 1$, with one exception in each case but with only $p > .10$. Clearly then, modified free recall after associative matching was not notably inflated under any circumstances.

The nonspecific-transfer group recalled fewer items under both lenient and stringent scoring [$F(1, 48) = 11.62$ and 7.22 , $ps < .005$ and $.025$ respectively]. The single-list and transfer group means, respectively, were 19.72 and 14.75 for lenient scoring and 17.06 and 12.12 for stringent scoring. In both analyses, these effects were independent of any interaction with retention-task sequence or direction of recall (W-R versus R-W), all $F_s < 1$. Recall in the W-R direction was significantly superior to R-W recall under both lenient and stringent scoring [$F(1, 48) = 28.34$ and 23.31 respectively, both $ps < .001$]. The interaction of direction of recall with number of lists learned

was—it should further be noted—also nonsignificant ($F < 1$).

For *associative matching*, retention-task sequence did not affect the subjects' ability to correctly match items [$F(1, 48) = 1.77, p > .10$; $X = 9.31$ for recall first versus 10.66 for matching first]. The subjects in the transfer group, however, again matched significantly fewer pairs than those in the single-list group [$\bar{X} = 8.90$ versus 11.10, $F(1, 48) = 4.94, p < .05$]. The main effect for list variants and all interaction effects yielded $F < 1$.

The correlations between errors on the recall task and correct scores on the subsequent matching task were also investigated. The mean number of errors on the recall task did not differ for the single-list and transfer groups [$t(62) = .50, p > .20$]. Hence, the correlations may be viewed as unaffected by differential error rates between the conditions. The product-moment r was $-.64$ for the single-list group and $-.23$ ($df\ 30$) for the transfer group ($p < .005$ and $> .05$). In agreement with the findings of Kanak and Curtis, therefore, the matching-task scores for the single-list group may have been adversely affected by prior recall-task errors. However, the lack of a similar, significant adverse effect for the transfer group implies that the present estimate of the retroactive inhibition of specific associations is conservative.

DISCUSSION

The results of the three experiments provided little evidence that the scores on associative matching were adversely affected by prior performance on any of the three modified free-recall tasks. Although Keppel (1966) reported such an effect with the mixed-list nondiscrete-trial recall procedure, the data from all three present experiments were consistent in demonstrating no differences in matching-task scores as a function of whether the matching followed or preceded the recall, for either the single-list or the nonspecific-transfer groups.

The correlations between recall-task errors and correct scores on the subsequent matching task could not be computed in Experiment II, which involved Keppel's recall procedure. In Experiments I and III the comparable correlations, when significant, were negative and were confined to the single-list group, which indicates, as Kanak and Curtis (1970) have suggested, that estimates of retroactive inhibition of specific associations for the nonspecific-transfer condition are likely to be conservative estimates relative to the single-list, control, condition, both under the unmixed-list unidirectional procedure and under

the discrete-trial bidirectional recall procedure.

The failure to replicate Keppel's results in Experiment II has no apparent explanation, but the outcome of Experiments I and III suggests that in any event Keppel's results may have little generalizability to recall procedures other than the mixed-list nondiscrete-trial technique. The present experiments reinforce the view that any confounding of matching-task scores by a prior recall task is likely to be limited to the single-list condition, and that as a consequence, estimates of the degree of retroactive inhibition in transfer paradigms are likely to be conservative.

Although Keppel also reported that the scores of *modified free recall* under lenient scoring were inflated after performance on a matching task, a direct attempt to replicate this effect in Experiment II was not successful—under either lenient or stringent scoring. In addition, this effect was not found for either type of scoring procedure in Experiment III, which used discrete-trial bidirectional recall, or under stringent scoring in Experiment I, which used unmixed-list unidirectional recall. Under lenient scoring in Experiment I, a statistical trend in support of this effect was observed, but it is important to note that the trend was confined to the transfer condition, where the prior matching task apparently reinstated the availability for recall of items that had been weakened through the competition between context-item associations of lists 1 and 2. Apparently the procedures of Experiments II and III, which required recall of both types of items (W and R items), introduced enough more difficulty than the procedure of Experiment I that subjects benefited little from prior observation of the to-be-recalled items.

Again, the discrepancy in Experiment II from Keppel's results, combined with the null effects of Experiments I and III for the single-list condition, suggests that the 'inflated-recall effect' which Keppel reported after associative matching has limited generality across recall procedures. The presence of such an effect, although admittedly weak, in Experiment I for the transfer condition indicates that giving the matching task first and employing unidirectional recall may be useful in the study of the reinstatement of the availability of 'unlearned' items in other retroactive-inhibition paradigms, both for verbal-discrimination and for paired-associate learning.

Prior evidence (Kanak and Curtis, 1970) on the unlearning of *context-item* associations in the nonspecific-transfer condition (relative to the single-list condition) was strengthened by all three experiments. This unlearning effect was statistically significant for both lenient and

stringent scoring in all analyses except for lenient scoring in Experiment II, where the difference still yielded a substantial trend. Although the present experiments did not employ a control condition for normal forgetting over time, Kanak and Curtis demonstrated no differences between an immediate-recall control group and the normal-forgetting control. Recall in the W-R direction was superior to R-W recall in Experiments I and III, but there was no difference in Experiment II. The latter results may be attributable to the mixed-list nature of the recall task or to a reliability-related effect, in that fewer W and R items were recalled by a given subject than in the other recall procedures. More importantly, however, context-item unlearning occurred for both types of items in each experiment.

Experiments II and III also confirmed the previously reported (Kanak and Curtis, 1970) unlearning of *specific* list 1 associations in the transfer condition. Although Experiment I failed to find this effect, the absolute difference was in the expected direction. As Kanak and Curtis and others have suggested, the phenomenon of unlearning of specific associations represents an interesting theoretical problem, since no task-specific mechanism appears responsible. It is conceivable, however, that the similarity of contextual cues at the time of matching-task performance to contextual cues at the time of original learning may serve to produce a functional competition mechanism in the matching task itself.

Notes

Dr. Cole is now at the Department of Psychology, Arkansas Polytechnic College, Russellville, Arkansas. Received for publication September 7, 1971.

References

- Briggs, G. E. 1954. Acquisition, extinction, and recovery functions in retroactive inhibition. *Journal of Experimental Psychology* 47:285-293.
- Kanak, N. J., and Curtis, C. D. 1970. Unlearning of context-item and specific associations in the W_1 - R_1 , W_2 - R_2 paradigm. *American Journal of Psychology* 83:256-263.
- Kausler, D. H., and Olson, R. D. 1969. Homonyms as items in verbal-discrimination learning and transfer. *Journal of Experimental Psychology* 82:136-142.
- Keppel, G. 1966. Association by contiguity: Role of response availability. *Journal of Experimental Psychology* 71:624-628.
- McGovern, J. B. 1968. Extinction of associations in four transfer paradigms. *Psychological Monographs* 78(whole no. 593).
- Palermo, D. S., and Jenkins, J. J. 1964. *Word association norms: Grade school through college*. Minneapolis: University of Minnesota Press.

TEMPORAL DATING AND SINGLE-TRIAL FREE RECALL

C. James Goodwin and Darryl Bruce

Florida State University

The temporal dating (knowledge of an item's position in a list) of subjects who had been told they would be tested by free recall was then assessed by a reconstruction task, after they had received either one list or six. Other subjects were tested by free recall after the same amounts of practice. Free recall showed a drop in primacy with practice whereas reconstruction did not, which suggests that temporal tags per se may be relatively unimportant as retrieval cues for the recall of early list members.

For present purposes, 'temporal dating' will refer to a subject's knowledge of an item's position within a list of items. Recently, this feature of memory has begun to attract increasing attention (Peterson, 1967; Tulving and Madigan, 1970; Underwood, 1969). Nevertheless, research on the topic has been relatively sparse, and Tulving and Madigan have pointed to the need for systematic study of temporal dating, especially its relations with other aspects of memory. The question in the present study was whether or not temporal knowledge has any functional value in free recall.

One easily demonstrable finding in single-trial free recall is a change in performance as a function of practice. The serial-position curve for early lists is characterized by a marked primacy effect and a somewhat smaller recency effect. After a number of lists have been presented, however, the curve shows more recency than primacy (Raymond, 1969). For reasons to be indicated presently, the emphasis in the present study was on the reduced primacy effect. The specific question was whether or not a similar change would occur if subjects expecting free recall were asked instead for knowledge of the order of presentation at two stages of practice.

The task used to assess temporal dating was *reconstruction*. There were a number of reasons this procedure was chosen. First, reconstructing the original order of a list can be thought of as analogous to making a series of recency judgments on pairs of items from a list (Tulving and Madigan, 1970). Hence it can be argued, at least logically, that temporal judgments are a fundamental part of a reconstruction

task. Second, the method is a very practicable one. Third, it has been reported (Underwood, 1969) that standard serial-position effects are obtained with this procedure—a necessity, of course, if one is interested in changes in such effects with practice. However, one particularly strong assumption in the present use of this method must be noted, namely, that reconstruction assesses the same temporal information that might be used in recall. At this time, there appear to be no firm empirical grounds to support (or deny) this assumption.

The main data of interest were serial-position curves. Curves for reconstruction indicating the same reduction in primacy that occurs with practice in single-trial free recall could be interpreted to mean that temporal tags are at least partly responsible for the recall of early items of a list. More generally, this might imply that the temporal date of an item can be an important retrieval cue (Tulving, 1968; Tulving and Patterson, 1969). Of course, the absence of any decline in reconstruction primacy with practice could imply just the opposite, namely, that temporal information is not very functional in free recall, at least for initial list members.

METHOD

—Materials and equipment—The items for each list were single- or double-syllable concrete nouns taken from the tables of Kučera and Francis (1967). The median frequency of the words selected was 39 per million. Responses to each test were written on a separate page in a booklet. A slide projector controlled by timer was used to present the lists.

—Procedure and experimental design—All subjects read the same instructions from typewritten sheets. They were told to expect a series of lists of words and that after each list they would be required to perform one of two tasks, free recall or reconstruction. They were to be told before presentation of each list which test would be administered. The time allowed for the free-recall task was 45 sec; the reconstruction task was unpaced and generally took between 60 and 90 sec. Presentation rate was one word per 1.5 sec. All groups received two four-item sample lists, the first tested by free recall and the second, by reconstruction.

The two independent variables of interest were degree of practice—subjects were either naive (N) or practiced (P)—and the type of test given after the final list—free recall (FR) or reconstruction (R). Both were between-subject variables and were combined factorially to yield four main conditions: NFR, PFR, NR, and PR. The N subjects were given one list, while the P subjects were given six. The final list given to each subject, either list 1 or list 6, was the one of major concern (the *critical* list).

The free-recall test, of course, involved recalling the list members in any order. In the reconstruction task, subjects were re-presented the list in a

scrambled order in their response booklets and asked to rewrite it in its original order of presentation. Inadvertently, only one randomized test order was used for all critical lists. However, there was no reason to suspect that this randomization was peculiar in any way. Moreover, a different randomization was used for all noncritical lists tested by the reconstruction procedure, yet the serial-position curves based on the two randomizations were virtually identical. In short, this restriction on the ordering of the reconstruction-test items would not seem to place any severe limit on the generality of the results to be reported.

Members of group NFR were given one free-recall list. The subjects in condition NR were given a 'misinstructed' reconstruction test. That is, before list presentation these subjects were instructed that they were to be tested by free recall, but after presentation, they were told to turn to the final page in their booklets and to complete the reconstruction test found there. This was done so that the critical list would be processed in the same manner for both NFR and NR groups.

Condition PFR was divided into two subgroups. Subgroup PFR-1 received six free-recall lists. Subgroup PFR-2 received two free-recall lists, one reconstruction list preceded by correct instructions, and then three free-recall lists. Condition PR also contained two subgroups. Subgroup PR-1 received five free-recall lists followed by one misinstructed reconstruction list. Subgroup PR-2 received two free-recall lists, one reconstruction list preceded by correct instructions, two more free-recall lists, and finally a misinstructed reconstruction list. In view of the initial instructions that both free-recall and reconstruction tests would be used, these subdivisions of the P (practiced) groups seemed advisable, since they permitted a determination of whether any peculiarities in performance occurred from the use of an uninterrupted sequence of free-recall lists prior to the critical list.

One set of six 12-item lists was used for all four conditions. In the N (naive) groups, each list was used equally often, and in the P groups, each occurred equally often at each position in the sequence of six lists.

For the R (reconstruction) groups, there was a lag of about 8 sec from the end of the critical-list presentation until the time the subjects had received the new instructions (the change from free recall to reconstruction) and had turned to the last page of their booklets. Consequently, an equivalent delay followed the critical list for subjects in the FR (free-recall) groups. This was accomplished by telling them to turn to the final page of their booklets to recall. The instruction was of equivalent length and duration as the instruction to the R groups. Since such a delay seemed likely to eliminate the recency effect of free recall, the focus of the present study was on differential primacy effects as a function of practice.

—Subjects—The subjects were 144 male and female students at Florida State University who served in fulfillment of a requirement of the course in introductory psychology. There were 36 subjects per condition and 18 in each PFR and PR subgroup. Up to three subjects were tested at a time. Their assignment to a particular condition was derived from a schedule drawn up from a table of random numbers. A complete cycle through all the conditions did not necessarily occur before the next cycle began, although the schedule did not depart too radically from this restriction.

RESULTS

Responses on the free-recall tests were scored correct even if misspelled or pluralized. For the reconstruction tests, subjects were given credit for a correct response if the output position of the particular item in question did not differ from its input position by more than one (see Zimmerman and Underwood, 1968). To analyze differences in serial-position curves, 'macrosubjects' were formed. In general, this was accomplished by summing over the subjects in each condition who were tested on the same critical list. Analyses were actually carried out on proportions to which the arcsin transformation had been applied. For descriptive purposes only, the serial-position curves to be presented have been smoothed somewhat by plotting a moving average consisting of each serial position as well as the two serial positions on either side of it. The first and twelfth positions were weighted by a factor of two and averaged with the second and eleventh positions respectively.

To begin with, it was found that there were no reliable differences between the serial-position curves on the critical lists for either PFR-1 and PFR-2 [$F(1, 10) < 1$] or PR-1 and PR-2 [$F(1, 10) < 1$]. On the basis of these analyses, the subgroups were combined to form the main groups PFR and PR for further analysis.

Primacy/recency shift for noncritical lists

Figure 1 indicates the effect of practice on the serial-position curve for single-trial free recall. Data from list 1 of groups PFR and PR were combined to form one curve, and data from list 5 of these groups were combined to form the other. This seemed reasonable in view of the absence of any significant differences between PFR and PR [$F(1, 10) < 1$] and any significant interaction involving this factor. For list 1 there was a pronounced primacy effect but little recency. By list 5, however, the primacy effect had dropped considerably and the recency effect had become greatly elevated. This shift was reflected in a significant interaction of lists and serial position [$F(11, 110) = 3.58, p < .01$].

Performance on critical lists

The left panel of Figure 2 demonstrates the effect of practice on the serial-position curves for the critical lists of the FR (free-recall) groups. The primacy portions of these curves resemble those in Figure

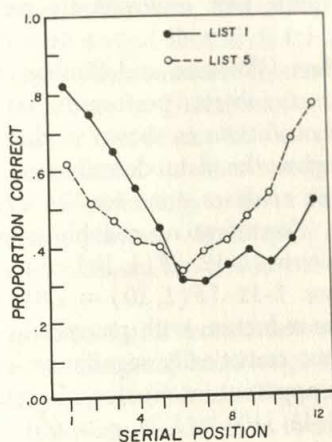


Fig. 1. Proportion correct as a function of serial position for the PFR and PR conditions combined on list 1 and on list 5

1; that is, there was again a marked drop in primacy with practice. Analysis of variance confirmed the substance of the difference between the two FR groups both for the entire list [$F(1, 10) = 19.94, p < .01$] and for serial positions 1-6 [$F(1, 10) = 16.90, p < .01$]. However, there was no reliable difference between the groups for positions 7-12

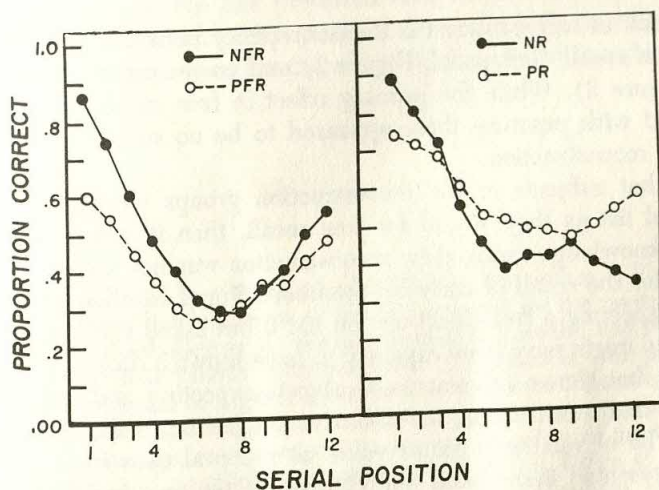


Fig. 2. Proportion correct as a function of serial position for free recall (left panel) and reconstruction (right panel) at two stages of practice

[$F(1, 10) < 1$]. This latter fact undoubtedly reflects the delay in testing the critical list, for it is well known that delaying free recall reduces the recency effect (Postman and Phillips, 1965).

The effect of practice on subjects performing on the critical list for the R (reconstruction) conditions is shown in the right panel of Figure 2. In order to compare the data described in the two panels of Figure 2, the same three analyses done for the FR groups were also done for the R groups. There was no reliable effect due to stage of practice for serial positions 1-12 [$F(1, 10) < 1$], positions 1-6 [$F(1, 10) < 1$], or positions 7-12 [$F(1, 10) = 1.64, p > .05$]. Further analysis showed that the reduction with practice in the primacy effect at positions 1-2 was not statistically significant either [$F(1, 10) = 1.68, p > .05$]. Also, the apparent interaction of groups and serial position was not reliable [$F(11, 110) = 1.34, p > .05$].

A direct statistical comparison of FR and R conditions for the entire list was not considered feasible because of the large amount of subject variance for reconstruction relative to free recall. However, a comparison was made over serial positions 1-6, where the variance difference between the two tasks was not as great. The interaction between type of task and stage of practice, however, was not significant [$F(1, 20) < 1$].

DISCUSSION

The result of major interest is the discrepancy between the effect of practice on recall (left panel, Figure 2) and on reconstruction (right panel, Figure 2). While the primacy effect in free recall was clearly diminished with practice, there appeared to be no such well-defined change in reconstruction.

Given that subjects in the reconstruction groups were processing the critical list as they would for free recall, then it seems that the temporal knowledge assessed by reconstruction was not a very prominent cue for the recall of early list members. For if so, then practiced subjects expecting a free-recall test on list 6 but asked for reconstruction instead might have been expected to have shown a drop in primacy similar to that shown by practiced subjects expecting and ultimately tested by free recall. Thus, the tentative suggestion is that temporal tags may be of relatively minor value as retrieval cues for recall of initially presented items—and that changes in the way in which temporal tags are stored and used may not be the principal cause of shifts in the recall-primacy effect as subjects acquire more practice.

Other interpretations of the data, of course, might be entertained. Consider first the fact that temporal information clearly seems available in both free recall and reconstruction. This is indicated both by the present results and by those of Underwood (1969) and Zimmerman and Underwood (1968). Now, it may be that this information is more accessible under reconstruction. In other words, perhaps in reconstruction the threshold for temporal tags is lower than in free recall. If so, then the fact that recall primacy significantly declined with practice whereas reconstruction primacy did not is just what would have been expected.

Alternatively, it could be that a reconstruction test is not an adequate measure of the type of temporal knowledge that subjects might use in the recall of initially presented items. That is, the temporal information used in the two tasks may be different in some way. Perhaps, then, serial recall would be a better procedure for assessing the functional value of temporal cues in free recall, the assumption being that the temporal information used in this task approximates that used in free recall more closely than do the time tags mediating reconstruction. Or finally, it could be that the large amount of subject variance in reconstruction simply masked any significant decrement in primacy that may occur with practice on this task.

In sum, the observations reported in this paper might well be considered susceptible to more than a single interpretation. But certainly one of them, and the one preferred here, is that the results are not particularly consistent with a point of view which stresses the efficacy of temporal tags as cues for the retrieval of early list members during free recall.

Notes

This paper is based on a thesis submitted by the first author to Florida State University in partial fulfillment of the requirements for the master's degree. The authors thank George E. Weaver, Daniel S. Lordahl, and Nicholas L. Rohrman for their various contributions to the research. The project was performed pursuant to a grant from the U.S. Office of Education, Department of Health, Education, and Welfare. The opinions expressed herein, however, do not necessarily reflect the position or policy of the U.S. Office of Education, and no official endorsement by the U.S. Office of Education should be inferred. Send reprint requests to C. James Goodwin, Department of Psychology, Florida State University, Tallahassee, Florida 32306. Received for publication August 20, 1971.

References

- Kučera, H., and Francis, W. N. 1967. *Computational analysis of present-day American English*. Providence: Brown University Press.
- Peterson, L. R. 1967. Search and judgment in memory. In *Concepts and the structure of memory*, ed. B. Kleinmuntz, pp. 153-180. New York: Wiley.
- Raymond, B. 1969. Short-term storage and long-term storage in free recall. *Journal of Verbal Learning and Verbal Behavior* 8:567-574.
- Tulving, E. 1968. Theoretical issues in free recall. In *Verbal behavior and general behavior theory*, ed. T. R. Dixon and D. L. Horton, pp. 2-36. Englewood Cliffs, N. J.: Prentice-Hall.
- Tulving, E., and Madigan, S. A. 1970. Memory and verbal learning. *Annual Review of Psychology* 21:437-484.
- Tulving, E., and Patterson, R. D. 1968. Functional units and retrieval processes in free recall. *Journal of Experimental Psychology* 77:239-248.
- Underwood, B. J. 1969. Attributes of memory. *Psychological Review* 70: 559-573.
- Zimmerman, J., and Underwood, B. J. 1968. Ordinal position knowledge within and across lists as a function of instructions in free recall learning. *Journal of General Psychology* 79:301-307.

BOOK REVIEWS

Julian Hochberg
Columbia University

Explanation in the Behavioural Sciences

Edited by Robert Borger and Frank Cioffi. New York: Cambridge University Press, 1970. Pp. xxi, 520. \$15.00.

The book consists of 12 invited papers on a variety of topics related to explanation in psychology, sociology, and anthropology: reasons versus causes as explanations (Toulmin), the explanation of purposive behavior (C. Taylor), the brain as a physical system (Sutherland), the uselessness of conditioning as an explanation (Hamlyn), rationality as a universal explanation of behavior (Watkins), cultural relativity and problems of explanation in sociology (Jarvie), individual behavior and the emergence of group properties (Wisdom), the relevance of psychology to the explanation of social phenomena (Homans), the Skinnerian analysis of behavior (Boakes and Halliday), explanation and the concept of personality (Eysenck), problems of explanation in linguistics (Chomsky), Freudian theory as pseudoscience (Cioffi). However, the format is novel in that following each essay is a critical commentary (the commentators are R. S. Peters, R. Borger, J. H. Grundy, A. J. Watson, A. Donagan, P. Winch, R. Brown, P. Blau, K. Pribram, D. Bannister, M. Black, and B. A. Farrell), followed in turn by a rebuttal.

The idea of a dialogue has, of course, a long philosophical tradition, but the device is less common in collections of scientific papers. It is helpful here: the papers are probably clearer than they would have been without the threat of adverse comment (although this is hard to judge); the commentators' occasional failures to understand an author's argument make it easier for the reader to distinguish deficiencies in himself from defects in what he is reading; and many commentators provide a useful summary of the essential points in the subject of their commentary.

I will confine my comments to three broad issues that emerge from a reading of this collection: the transformational approach to understanding linguistic behavior, represented by Noam Chomsky and, commenting on his paper, by Max Black; the notions of rationality and reinforcement, represented by the papers and comments of Watkins, Toulmin, Donagan, Taylor, Borger, Boakes and Halliday, and Pribram; and the problem of reductionism, represented by Sutherland, Grundy, Wisdom, Brown, Homans, and Blau.

Chomsky's essay displays his usual polemical talents, laced with exaggera-

tion ("behaviorist approaches to language and mind are grounded in a faith in the shallowness of explanations, a belief that the mind must be simpler in its structure than any known physical organ"; p. 426), with linguistic insight, and with obscure scholarship (Chomsky's familiarity with the works of the Russian poet V. Shklovskij is unlikely to be shared by many of his readers). It is also laced with a rather broad unawareness of much of the behavioristic tradition: according to Chomsky, "the most serious defect of classical philosophy of mind, both rationalist and empiricist, seems to me its unquestioned assumption that the properties and content of the mind are accessible to introspection; it is surprising to see how rarely this assumption has been challenged" (p. 426). Surprising to Chomsky perhaps, but not to readers of Watson, Freud, and their intellectual descendants.

Eminence is entitled to its eccentricities, and Chomsky's scholarly defects at least have the merit of enlivening his writings. Nevertheless, it is a little startling to see accusations about simpleness of mind directed solely at behaviorists, since no matter how guilty they may be (and who doesn't have a favorite behavioristic Aunt Sally?), very similar views are shared by others who cannot easily be tarred with the behavioristic brush. Herbert Simon, for example, has frequently expressed his belief that the complexity of human behavior reflects not the complexity of mind but the complexity of the human environment. The picture of Skinner and Simon on one side, as advocates of the essential simplicity of man, arrayed against Chomsky on the other, as advocate of his complexity, is likely to offend all three.

Polemics aside, Chomsky's paper and Max Black's comments on it raise a number of fundamental questions about the transformational approach to understanding linguistic behavior. The central problem, often remarked, is the distinction between grammatical and other (semantic, phonological) determinants of verbal behavior. The notion of 'grammaticalness' can be considered as an instance of the old problem of universals: what is the 'essence' of a dog, a table, a triangle, or a grammatical sentence? In most cases we can *recognize* instances of each, but grave difficulties are encountered when an attempt is made to set down a general definition that corresponds to our judgments. Judgments of 'grammaticalness' are perhaps more difficult to make—there are more borderline cases, and dependence on context is greater—than in the case of the other universals just mentioned.

Moreover, in many of the more familiar cases, some general definitions can be given of a functional ('flat surface raised above the ground on which objects may be placed'), geometrical ('three-sided figure'), or biological (species) kind, definitions that are tolerably successful. Presumably the search for a transformational grammar for English (or more ambitiously, for a 'universal' grammar) resembles the geometrical case: with the 'idea,' intuitively given, of a set of entities (triangles, grammatical sentences), the task is to find the set of formal rules that allows us uniquely to generate those entities. In the case of triangles, 'three-sided figures' (with appropriate qualifiers) does the job, and the existence of this definition validates the concept. A transformational grammar for English is yet to be found, so that the usefulness of this idea is yet to be demonstrated.

Chomsky's self-imposed task is often considered (and he himself appears to consider it) as a scientific one: to explain one aspect of linguistic behavior.

(I accept Black's comment that much of Chomsky's emphasis on 'mind,' 'rationalism,' and so on, is *pour épater les empiristes*, so I don't feel I am doing him an injustice by using behavioral language.) Judged in these terms, there are serious questions to be raised about his approach. As Black points out, natural speech does not put one in mind of a precise mechanism, operating according to a limited set of axiomatic rules. There are slips of the tongue, inconsistent grammatical errors (for example, the use of different grammatical forms in different social contexts), hesitations, memory lapses, and so on. To avoid these obvious difficulties, the distinction between linguistic *competence* (what the speaker knows about grammar) and linguistic performance (what he says) has become generally accepted among psycholinguists. This is a familiar kind of distinction: it resembles the psychological dichotomy between 'learning' and 'performance,' and the biological one between 'genotype' and 'phenotype.' In each case, it is possible (although I believe this to be an error) to think of the dichotomy as one between an underlying reality and an external manifestation that is an imperfect copy of that reality; this is Plato's parable of the cave.

Consider first the genotype/phenotype distinction, the best understood of the three. In what sense is the phenotype a more or less perfect 'reflection' of the genotype? In no sense at all, in fact. Phenotypic characters are not contained as perfect (universal?), preformed prototypes in the genome, and every phenotype is as intimately related to its genotype as every other: no phenotype is more 'perfect' (or 'normal' in the normative sense) than any other, since all must develop epigenetically under the joint, reciprocal influence of the genotypic instructions and the environment within which they work. The hereditary material simply works under the computerman's principle Garbage In Equals Garbage Out, except that the output is partially incorporated into the structure of the developing organism. Yet everyone, save professional biologists and others who have devoted considerable thought to the matter, acts as if this were not true: 'obviously,' the shrunken, deformed phenotype that has developed under conditions of nutritional deficiency is less normal (grammatical?) than the symmetrical, well-formed phenotype that is the product of healthy rearing conditions. We are all Platonists at heart, and it is only with considerable effort that this tendency can be overcome, even when the generative mechanism is quite well understood. This tendency is probably one source of the confusion that continues to surround the issue of the genetic determination of IQ. When, as in the case of language, the generative process is not understood, the difficulties in seeing beyond simple characterizations like 'grammatical' versus 'non-grammatical' or 'meaningful' versus 'nonsensical,' the difficulties of seeing beyond them to the nature of the underlying mechanism (or mechanisms)—and thus to the relationship between it and the stream of speech as it is observed *in vivo*—are immensely magnified.

As I have pointed out elsewhere, these strictures are not restricted to the genotype/phenotype distinction; there are many reasons to suppose that the relationship between behavior ('performance') and its underlying generative mechanisms ('mechanisms of behavioral variation') is every bit as complex and epigenetic. Hence it is not obvious that an *insistence* on the *a priori* validity of categories, such as syntactic or semantic, is likely to be

helpful in unraveling the mechanisms of linguistic behavior. If geneticists and embryologists had restricted their attention to 'normal' organisms, genetics would have been greatly retarded. Claude Bernard has rightly remarked that the exceptions are as important as the normal: the same laws must apply to all. So Chomsky's predilection for studying what the speaker knows about what he says, rather than what he says, must be regarded with suspicion—at least if the aim is to understand human behavior.

To what extent is it permissible for the theorist to demarcate his subject matter *before* he explains it? Just because people are capable, albeit imperfectly, of distinguishing between grammatical and nongrammatical sentences, independently of meaning, context, and so on, is not to say that the distinction has any special validity at the level of the generating mechanisms. We can recognize triangles as opposed to other objects, independently of their color, size, location, and so on; yet few would maintain that this implies the existence of some Platonic triangle-recognizing system that can be studied independently of the process of space and pattern recognition as a whole. Explanations are where you find them, and there is little in the history of science (especially biological science) to support the idea that the theoretical endeavor is aided by the kind of a priori dichotomies current in psycholinguistics. Hence, the distinction between competence and performance may be quite spurious when it is used artificially to limit the domain of theories about language. It is true that people know more than they say; but it is questionable whether what they know can be studied independently of what they say, or whether either can be explained independently of the other.

Perhaps the key to the paradox presented by Chomsky's ideas is that his aim is not really a scientific one. Certainly he is not sympathetic to an evolutionary, hence a biological, epistemology, even though his subject matter is biological. For example, he disagrees with Lorenz's contention that evolution by natural selection has any implications for the Kantian framework within which Lorenz would place the findings of comparative ethology. This is natural since, as Mayr has often pointed out, a Platonic, 'essentialist' metaphysic is antithetical to the idea of evolution.

Chomsky does not have a firm grasp of the epigenetic nature of behavioral development, as indicated by his confident use of the term 'innate.' To elaborate on what 'innate' might mean when applied to language universals (if such there be) would take us well beyond the confines of a review. At the very least, the term requires more defense than Chomsky chooses to muster.

My conclusion is that Chomsky's approach represents a fusion of two distinct elements. The first may, with only slight exaggeration, be termed Pythagorean: his evident conviction that language, at least in its grammatical aspects, is fundamentally mathematical. This conviction, directed toward physical nature, animated many of the founders of modern science, but it has been much less successful in biology. The second element may be termed Boolean, after George Boole, the nineteenth-century English mathematician who discovered mathematical logic but thought he had found the Laws of Thought. Like Boole, Chomsky seeks to formalize an aspect of human behavior, defined a priori; unlike Boole, he has not yet succeeded. But, like

Boole, it may be that even if he succeeds, his achievement will be regarded as a mathematical rather than a psychological one. People and animals are sometimes logical and sometimes illogical, and people are sometimes grammatical and sometimes not; but no matter how much we learn about either logic or grammar, it is likely that the problem of understanding human behavior—rational and irrational, grammatical and ungrammatical—will remain.

The persistent attempt to understand human behavior in terms of the idea of rationality has many similarities to Chomsky's attempt to mathematize syntax. It differs, however, in that to explain behavior in terms of objectives and rational means to attain them is to espouse (in most cases) a teleological rather than a mechanistic form of explanation. A generative grammar allows us to produce language, but a rationalistic explanation of an action is post hoc; it tells us why someone did something, but not the means whereby he arrived at his decision.

Thus, Watkins describes Hough's ingenious explanation of the apparently bizarre behavior of the respected nineteenth-century British admiral, Tryon, who apparently ordered the ships of his fleet to collide with one another, which they duly did, killing him in the process. Hough explains this incident by means of a set of plausible motives and actions derived from them. Yet these motives—to turn the fleet, to keep his officers on their toes—did not require the daring and, as it turned out, ambiguous series of orders that Admiral Tryon gave. To infer from his motives the particular actions that he took requires more than a knowledge of the motives; it requires a knowledge of Tryon. And indeed this reconstruction was first arrived at by a friend of Tryon's.

To take another example: a man may wish to exercise for the sake of his health; however, his choice of swimming over golf, tennis, or some other sport may not be explicable in this way, since all may be equally satisfactory solutions to the health problem. There are two kinds of solution that can be offered here. One is to preserve the 'rationality principle' by postulating additional motives—a preference for noncompetitive versus competitive sports, for example. The other, which I would favor, is to recognize that whether or not a man's behavior can be understood in terms of his aims, his ability to attain those aims depends on his nature—his resourcefulness, the strategies he has learned or can devise, and so on. If these restrictions are not understood, both apparent irrationality (such as Admiral Tryon's puzzling actions) and compulsions (such as overeating) strain the notion of rationality as a general explanation.

This alternative is similar, but not identical, to Watkins' contention that an individual's unsuccessful actions should be explained in the same way as his successful ones, that is, in terms of his objectives and rational means to obtain them. I agree with this 'uniformity principle,' but I believe that if it is applied broadly, to all behavior, including things like compulsions and impulsive acts, it will require us to understand the *mechanisms* underlying human action, and not just people's motives.

This latter distinction is difficult to communicate when the complexities of human behavior are at issue, but as Wiener was perhaps the first to point out, it is easier to see in the case of a simple servomechanism. Thus, the

'motive' of a radar guidance system may be said to be to keep its missile on target. However, the mechanism by which it attains this objective is the whole complex of sensors and feedback loops that tracks the missile, compares its course with the desired one, and sends out appropriate corrective signals. With a knowledge of motive alone, we are able to understand why the missile hits the target, even when its course is perturbed; but with a knowledge of mechanism, we are in a position to understand both the conditions under which it hits the target and those under which it does not.

All these arguments apply with little alteration to the concept of reinforcement, which as it is usually used, is (like motive) teleological. Like the guidance system, reinforcers achieve their effects via mechanisms within the organism. Like advocates of the rationality principle, thoroughgoing reinforcement theorists are prone to postulate additional reinforcers (motives) if the observed behavior does not make sense in terms of the reinforcers that are known to be operating. The *reductio ad absurdum* here is to postulate that activities themselves (rather than the opportunity to engage in certain activities, which is a more legitimate notion) may be reinforcing. In this way, embarrassing findings such as that hungry pigeons continue to peck a light even when pecks prevent the delivery of food can be explained: pecking is clearly a self-reinforcing behavior under these conditions. A more defensible strategy here is to consider a behavior such as pecking and the mechanisms that must underly it and to ask what the nature of these mechanisms must be that they produce adaptive pecking in one set of situations, no pecking at all in others, and maladaptive pecking in still others. This is simply an application to the problem of reinforcement of the uniformity principle just enunciated.

It is puzzling that the problem of reductionism is still with us. It appears at all levels. In psychology it involves the relationships between physiology and behavior—can the properties of behavior be deduced from the properties of neurons? In sociology it involves the relationships between individuals (and individual psychology) and the group (and group sociology) of which they are a part—are there emergent 'group properties' that cannot be reduced to the behavior of individuals? The persistence of this issue is puzzling because it seems so obviously to be a matter of the stage of development of the science in question. For example, we now have a reasonably good understanding of atoms, both as configurations of protons and electrons and as elements that can combine with one another in certain ways to form chemical compounds. The wave-equation description of intra-atomic structure can, at great cost in terms of convenience, be extended to describe the domain of chemical compounds. Chemical laws remain more convenient for chemical phenomena, but there is no longer any mystery about the epistemological relationship between atomic and chemical levels of explanation.

However, there was a time when the intra-atomic level, represented by experiments on X-rays, alpha particles, and so on could not be related to the chemical level at all; still earlier, only the chemical level was known. It seems clear, therefore, that the problem of reductionism is really a problem of incomplete knowledge. One asks whether behavior can possibly be explained by interactions between neurons only because no one has yet shown how the two domains are related. Philosophers may object that the problem

of consciousness, for example, cannot yield to a reductionistic account, but I believe that they underestimate the power of a good explanation to disperse (perhaps only after several generations) objections of this sort. Philosophical contemporaries of Newton, for example, must have thought his failure to deal with the problem of action at a distance a critical flaw in his theory, but no one worries about it any more.

Few mature experimental scientists are preoccupied with the philosophical issues discussed in this collection. Most people have either solved these problems to their own satisfaction by the time they reach mid-career or, more frequently, have decided that excessive concern with issues of this sort is a kind of lotus-eating that is inimical to robust mental health and professional success alike. However, good science, like good statecraft, is the outcome of a proper balance between action and reflection, although, as Edison's apothegm reminds us, the balance is usually thought to favor action. On the off chance that action is perhaps overrepresented, I welcome this collection as a stimulus to thought. The major flaw of the book is its lack of a subject index, which is particularly important in a collection of this kind; a comprehensive index would have done much to help the reader juxtapose differing approaches to the same topic by different authors. Also the editing is sloppy in places: the protean creation "Rescorla Solomon" (p. 352), for example, probably exaggerates the degree of collaboration between Rescorla and Solomon in their 1967 paper; Sutherland's interesting article has no references and tantalizes the reader who would like to follow up some of his allusions.

J. E. R. Staddon, *Duke University*

Attention: Contemporary Theory and Analysis

Edited by David I. Mostofsky. New York: Appleton-Century-Crofts, 1970. Pp. vii, 447. \$12.95.

The collection of contributed papers *Attention: Contemporary Theory and Analysis* gives the impression of being a hardback issue of a journal devoted to the topic of attention. This impression is further strengthened by the fact that of the 18 papers included, 8 of them are experimental reports of original data not published elsewhere. The papers are loosely grouped into three sections entitled "Systematic Issues," "Perceptual and Learning Issues," and "Neuropsychological Issues." The fact that little attempt has been made to integrate these sections, or the papers within them, is best illustrated by the fact that there are no introductory or summary remarks for the individual sections, or for the book as a whole.

Any collection such as this is bound to be somewhat uneven in quality. There are a few very fine papers in this collection and a few quite weak. Among the former is an interesting theoretical contribution by Hochberg, in which he advances the novel proposition that selective attention is dependent on the testing of expectancies about the organization of the incoming stimulus array. There is also an excellent analysis of multichannel-processing models of perception, by Lindsay. As a methodological discussion, the use of the vigilance paradigm in the investigation of attention is treated with

some sensitivity by Jerison. There are also a few good reviews. Silverman provides a massive if somewhat rambling treatment of a broad gamut of issues, while Thompson and Bettinger competently handle the neural and electrophysiological data.

The major difficulty with any such treatment of attention is the lack of consistency in the very definition of the word 'attention.' Some of the authors interpret the concept as arousal, others as responsiveness to specific stimuli or as exploratory responses such as the orienting reflex, or even in terms of the amount of information input from the external receptors. The problem of definition becomes even more pronounced when one considers the interrelationship of attention and complex processes such as learning or perception.

It is clearly the case that if an organism learns a discrimination between two stimuli that differ along some dimension, we can take this as evidence the organism is attending to the stimulus dimension being varied. The generalization paradigm used in several of these papers provides less clear evidence for the operation of attention. If a subject responds to a new stimulus that differs slightly from the stimulus he has been trained on, is this evidence he is attending to the stimulus dimension in question and thus perceives the similarity between training and test stimuli, or is it evidence he is attending insufficiently to this stimulus dimension and thus does not detect the difference between test and training stimuli? Even in the simple discrimination task, we are on very thin ice when we say that an organism must attend to a stimulus in order to respond to it. How do we deal with the Soviet interoceptive-conditioning paradigm, where a subject learns a conditioned discrimination to stimuli that do not enter conscious awareness, such as the nature of liquid in the intestine? Must we say that although he is not conscious of these stimuli, he is nonetheless attending to them?

Such musings lead one to conclude that any definition of attention that does not include the subjective experience of the organism is probably inadequate. Perhaps it would have been better to attempt an organization around a specific definition or viewpoint of attention, or at least to attempt some sort of synthesis. Without such a clear definition and orientation, one must beware the warning of William James that "Attention may have to go, like many a faculty once deemed essential, like many a verbal phantom, like many an idol of the tribe. It may be an excrescence on psychology."

Stanley Coren, *New School for Social Research*

Human Perception

By R. H. Day. New York: Wiley, 1970. Pp. viii, 193. \$4.25.

The Intelligent Eye

By R. L. Gregory. New York: McGraw-Hill, 1970. Pp. 191. \$2.95.

Perception: Mechanisms and Models

Edited by Richard Held and Whitman Richards. San Francisco: W. H. Freeman, 1972. Pp. 390. \$5.95.

Three new books designed for use in undergraduate courses in perception have recently appeared. The most traditional in content and format is *Human*

Perception by R. H. Day. The positive attributes of this book include Day's usually clear writing and an extremely broad scope of coverage. The first two chapters deal with sensory transduction, methods of measurement, and basic sensory capacities. The ensuing chapters deal with the constancies, motion perception, brightness, color, the effects of learning, and adaptation to distorted inputs.

The best of the chapters are those dealing with the areas in which Day himself has done research. These are the sections on constancy (and its application to illusions), motion perception, and perceptual modification, which are clearly presented albeit theoretically rather one-sided. The most unfortunate aspect of this book is that it attempts to accomplish too much. In 183 pages it endeavors not only to cover all of the topics listed above but to incorporate data from vision, audition, and the chemical and mechanical senses as well. This results in rapid transitions from topic to topic, which makes the reading uncomfortably spasmodic. Frequently the material is spread too thin, with only mentions of topic areas rather than discussions. The text would have been considerably improved if some breadth had been sacrificed for more depth.

R. L. Gregory's book *The Intelligent Eye* seems at first glance to be rather attractive as a text, because of its numerous illustrations and dichromatic stereograms. On closer inspection, however, its appeal is sorely diminished. The book is based upon a series of BBC television programs and still bears most of the marks of a television special. An initial premise is adopted, namely, that perception is basically a judgmental process with a massive dose of learned patterns of cue utilization. The book then proceeds to offer only the data that support this proposition.

Some demonstrations are offered, frequently based on the numerous stereograms. Many of these stereograms are confusing because of the lack of sufficient introductory material. There is also a technical difficulty, because the green filter in the stereo spectacles is not well matched to the green ink in the stereograms. This leaves the reader with an unavoidable double image. There are numerous digressions, some of them interesting but most of them irrelevant to the task of presenting a background for the understanding of visual perception. A discussion of why the pyramids are shaped as they are, or of the Egyptian numerical system, may be entertaining but should not be favored over presentation of basic data on visual perception. I'm afraid that since the series was designed to maintain the interest of the lay television viewer, an orientation that is preserved in the book, this volume is better suited to the reader of popular science than to the serious student of perception.

The last book in this triad is a collection of readings from *Scientific American* that has been edited and introduced by R. Held and W. Richards and entitled *Perception: Mechanisms and Models*. Held and Richards have carefully selected a rather broad spectrum of important and interesting articles covering the basic sensory structures, physiological analyzers, illusions, constancies, basic perceptual organizational processes, development, and perceptual plasticity. Most of the articles are well written and relevant to the content of contemporary courses in perception. The book is lavishly produced with the usual excellent illustrative material that characterizes

Scientific American articles. There are a number of full color and innumerable bicolor plates, which hold the attention and illustrate major points in the papers.

The major feature in this collection, however, is the set of introductions by editors Held and Richards. These introductions provide a basic orientation for the student and attempt an integration of the material within each section. Several of the introductions are themselves illustrated, and almost all of them are helpful. The selection of articles combined with the introductory material make this book a useful accompaniment to most undergraduate courses in perception.

Stanley Coren, *New School for Social Research*

Behavior of Nonhuman Primates: Modern Research Trends, Vols. 3 and 4

Edited by Allan M. Schrier and Fred Stollnitz. New York: Academic Press, 1971. Pp. xv, 206; xv, 239. \$9.50, \$10.00.

The editors point out that at the completion of the first two volumes in 1965, no plans to continue the series were at hand. But with the wide acceptance of the initial effort and the dramatic increase in primate research, an extension of the series was warranted. The third volume continues in the spirit of the first two, authors with specialized interests and expertise summarizing the literature to date and bringing in their own recent data and communications in the field.

This particular effort, the third volume, might best be subtitled "Social Behavior and Psychophysics." Chapter 1, by Hinde on the development of social behavior, is an excellent up-to-date review of field and laboratory literature, with emphasis on his own latest work. Bernstein brings in new data in chapter 2 to combine with data already published to present a comprehensive picture of activity profiles of six primate taxa. His systematic analysis of behavior confirms many of my own suspicions from observations at zoos: that there are species-specific activity patterns that are distinct, indicative of phylogeny, and represent ecological adaptations.

Chapter 3, by DeValois and Jacobs is a review of the current status of primate vision research, and it reveals that the primate eye is very similar to the human eye and can function as a good model. The final chapter on hearing is a good example of why the book was needed. Most of the research in the field, except for some classics, has been done since 1965. This chapter by Stebbins is particularly strong on methodology and indicates that the field is no longer open to well-meaning amateurs.

All of the authors in the third volume have made significant contributions to the contemporary field of primate research. This is another good effort for each of them.

The fourth volume is quite different from the earlier volumes of this series. Whereas the first three volumes provided excellent summaries of the field and brought together relevant material already available in the literature, about half of the fourth volume provides us with new material that was relatively unavailable to individuals not 'on the inside' of contemporary

primate research. Actually, with the exception of chapter 1, this book is the study of three chimpanzees (Vicki, Washoe, and Sarah) who received a lot of attention and training during their developmental years.

Chapter 1 in this fourth volume is an extended analysis of an unpublished dissertation done at the University of Wisconsin. This chapter brings together, in great detail, information on retention of discrimination learning, with the emphasis on the learning-set paradigm. The chapter has the touch of Stollnitz's ability to be precise and inclusive when dealing with a great deal of material. The treatment of the data throws some new light on what is going on in the learning-set situation.

In chapter 2, Nissen and Nissen present the data collected on Vicki between 1947 and 1954. The data have not been reported in writing elsewhere, even though some of the information about this super-chimp was known to those who, like myself, worked at a primate laboratory for several years and met people who had had contact with the Nissens. This is an interesting chapter but would have been much better if the Nissens had updated and related their findings to contemporary research with children and nonhuman primates in the bibliography. As it is, the chapter seems to have been written in 1954 and the authors finally prodded into getting it into press.

The chapter on two-way communication with Washoe, by the Gardeners, goes far beyond the details of the 1969 *Science* paper. It is well written, very readable, and demonstrates in great detail what can be done when scientists utilize behavioral mechanisms that have ecological validity for the organism.

The last chapter, by Premack, on language competence in Sarah overlaps considerably with his previously published material as far as this reviewer can tell, but I might be missing something. He gives detailed procedures of the successive approximations used to teach Sarah the language. This is a fine demonstration that Skinnerian techniques can work, outside of a Skinner box, with a highly complex organism learning a very complex task. However, the author tends to ignore the rather non-stimulus/response transformations that Sarah is capable of making. I am also not sure of his conclusion that, because he had to teach Sarah rather than vice versa, language is unique to man.

All in all, this fourth volume is a book that most primatologists and comparative psychologists should have on their bookshelves for reference. I have found it quite useful already, since there were as many examples of ingenuity in animals reported in this book as have been reported in the last 20 years in the *Psychological Abstracts*. It also is a demonstration of how much we can learn from three animals.

Robert R. Zimmerman, *University of Montana*

Stimulus Properties of Drugs

Edited by Travis Thompson and Roy Pickens. New York: Appleton-Century-Crofts, 1971. Pp. xii, 211. \$13.50.

The stated purpose of this book was to "bring together investigators working in the field, and to review their own and other findings concerning the

possible stimulus functions drugs may serve" (p. xi). In actuality, few of the chapters do little more than review the work of the authors. (Two chapters are condensed doctoral dissertations and at least four more are reiterated versions of several of the authors' recent papers.) There is spotty coverage of the recent English-speaking literature and little-to-no coverage of the vast Eastern European literature in the field.

After the introduction, the book is divided into three sections, in which drugs are considered as unconditioned, discriminative, and reinforcing stimuli respectively. Three chapters, each followed by the comments of two discussants, are included in each section. The chapters tend to be very narrow in scope and often bog down in trivial discussions. One is amazed at the number of times various contributors marvel at the 'discovery' that drugs are actually like other types of stimuli.

On the positive side, two of the chapters are excellent. One is a review of the discriminative control of behavior by drugs (D. A. Overton) and the other a review of variables influencing drug self-selection (T. Thompson, G. Bigelow, and R. Pickens). Further, it is handy to those interested in behavioral pharmacology to have the major work of so many authors summarized in one volume. On the negative side, there is no author index, and the subject index is inadequate. One could not, for example, tell which chapters discuss a particular drug unless it happened to be a drug used in the research being presented. In summary, the book will be of some (if limited) use to people working in the field, particularly if they want a review of the research in the specific areas covered.

Stephen C. Woods, *Columbia University*

Books received

(The books listed here have not yet been noted in our pages. Listing here does not preclude their later review, however.)

- Alderfer, Clayton P. *Existence, Relatedness, and Growth: Human Needs in Organizational Settings*. New York: The Free Press, 1972. Pp. x, 198. \$7.95.
- Anthony, Sylvia. *The Discovery of Death in Childhood and After*. New York: Basic Books, 1972. Pp. 280. \$6.95.
- Ardile, Ruben. *Psicologia experimental: Manual de laboratorio*. Mexico City: Editorial Trillas, 1971. Pp. 140.
- Arnoult, Malcolm D. *Fundamentals of Scientific Method in Psychology*. Dubuque, Iowa: Wm. C. Brown, 1972. Pp. xvii, 217. Paperback, \$3.95.
- Aronson, Elliot. *The Social Animal*. San Francisco: W. H. Freeman, 1972. Pp. xiv, 324. Paperback, \$3.50.
- Aronson, Lester R.; Tobach, Ethel; Rosenblatt, Jay S.; and Lehrman, Daniel S., eds. *Selected Writings of T. C. Schneirla*. San Francisco: W. H. Freeman, 1972. Pp. xiv, 1032. \$22.50.

- Bannister, D., and Fransella, Fay. *Inquiring Man: The Theory of Personal Constructs*. Baltimore: Penguin Books, 1972. Pp. 221. Paperback, \$2.25.
- Barr, Murray L. *The Human Nervous System*. Hagerstown, Md.: American Medical Book Publishers, 1972. Pp. viii, 405. \$17.50.
- Barten, Harvey H., and Bellak, Leopold, eds. *Progress in Community Mental Health*. Vol. 2. New York: Grune & Stratton, 1972. Pp. xi, 290. \$9.75.
- Bastide, Roger. *The Sociology of Mental Disorder*, trans. Jean McNeil. New York: David McKay, 1972. Pp. viii, 251.
- Bean, Orson. *Me and the Orgone: The True Story of One Man's Sexual Awakening*. New York: Fawcett, 1972. Pp. 160. \$.95.
- Beidler, Lloyd M., ed. *Handbook of Sensory Physiology*. Vol. 4, *Chemical Senses 2: Olfaction*. Vol. 5, *Chemical Senses 2: Taste*. New York: Springer-Verlag, 1972. Pp. 518, 410. \$46.90, \$44.90.
- Biegert, J.; Leutenegger, W.; and Kummer, H., eds. *Proceedings of the Third International Congress of Primatology*. Vol. 1, *Taxonomy, Anatomy, Reproduction*. Vol. 2, *Neurobiology, Immunology, Cytology*. Vol. 3, *Behavior*. Basel: S. Karger, 1972. Pp. xiv, 278; x, 246; x, 191. \$22.80, \$21.60, \$16.80.
- Black, Abraham, H., and Prokasy, William F. *Classical Conditioning II: Current Theory and Research*. New York: Appleton-Century-Crofts, 1972. Pp. xii, 497. \$14.95.
- Boudon, Philippe. *Lived-in-Architecture*. Cambridge, Mass.: The M.I.T. Press, 1972. Pp. 200. \$7.95.
- Boyle, D. G. *Language and Thinking in Human Development*. London: Hutchison University Library, 1971. Pp. 191. £2; paperback, 90d.
- Brown, Alan R., ed. *Prejudice in Children*. Springfield, Ill.: Charles C Thomas, 1972. Pp. x, 214. \$6.95.
- Brozek, Josef, and Slobin, Dan I. *Psychology in the USSR: An Historical Perspective*. White Plains, N.Y.: International Arts and Sciences Press, 1972. Pp. x, 301. \$15.00.
- Cajal, Santiago Ramon Y. *The Structure of the Retina*. Springfield, Ill.: Charles C Thomas, 1972. Pp. xxxix, 196. \$12.50.
- Carek, Donald J. *Principles of Child Psychotherapy*. Springfield, Ill.: Charles C Thomas, 1972. Pp. xi, 226. \$10.75.
- Carroll, John B.; Davies, Peter; and Richman, Barry. *Word Frequency Book*. Boston: Houghton Mifflin, 1972. Pp. liv, 856.
- Carswell, E. A., and Rommetveit, Ragnar, eds. *Social Contexts of Messages*. New York: Academic Press, 1971. Pp. ix, 163. \$7.95.
- Cermak, Laird S. *Human Memory: Research and Theory*. New York: Ronald Press, 1972. Pp. viii, 294.
- Chein, Isidor. *The Science of Behavior and the Image of Man*. New York: Basic Books, 1972. Pp. xii, 347. \$12.50.
- Christiansen, Harley D. *Ethics in Counseling*. Tucson: University of Arizona Press, 1972. Pp. xi, 260. Paperback, \$5.95.
- Clark, Frank W.; Evans R.; and Hamerlynck, Leo A., eds. *Implementing Behavioral Programs for Schools and Clinics: Proceedings of the Third Banff International Conference on Behavior and Modification*. Champaign, Ill.: Research Press, 1972. Pp. xiii, 201. \$4.00.

New

Ernst Mach

His Life, Work, and Influence
John T. Blackmore

A wide ranging biography of the Austrian physicist and philosopher who had personal or intellectual ties to many extremely diverse major figures, from Planck to Lenin and from Einstein to Wittgenstein. Dr. Blackmore's dual abilities as philosopher and science historian enable him to integrate the relevance of various philosophic theories in the development of Mach's thought in physics, educational theory, psychology, and the history and philosophy of science.

LC: 79-138514 410 pages 22 halftones 7 line drawings \$16.95



from California

University of California Press • Berkeley 94720



Contributors to Volume 85

(Book reviews, notes of meetings italicized)

- Barker, Douglas, 431
 Beck, Jacob, 1
 Behar, Isaac, 129
 Bevan, William, 31
 Bitterman, M. E., 301
 Blackstone, W. T., after 462
 Braine, Lila Ghent, 157
 Bruce, Darryl, 597
 Buffardi, Louis, 377
 Cogan, Rosemary, 569
 Cole, Lawrence E., 57, 585
 Collyer, Stanley C., 31
 Coren, Stanley, 457, 611, 612
 Curtis, C. Dwayne, 585
 DaPolito, Frank, 431
 Deich, Ruth F., 393
 Duncan, Carl P., after 462
 Elmes, David G., 237
 Epstein, William, 21
 Erickson, James R., 249
 Evans, Rand B., 463
 Federspiel, Charles F., 99
 Fehr, Fred S., 63
 Flanagan, John L., 69
 Geldard, Frank A., 291
 Geller, E. Scott, 397
 Gergen, Kenneth J., 308, 458
 Gogel, Walter C., 201, 477
 Goldstein, Alvin G., 569
 Goodwin, C. James, 597
 Graham, Charles, 271
 Greener, William I., 237
 Haaseth, Kjell, 549
 Hall, James W., 277
 Hirsch, Joy, 241
 Humphreys, Michael S., 189
 Johannsen, Dorothea E., 308
 Johnson, G. J., 325
 Jones, Mari R., 249
 Jonides, John, 31
 Kanak, N. Jack, 57, 585
 Kasschau, Richard A., 43
 Kaufman, Lloyd, 141
 Kendon, Adam, 441
 Kurtz, Kenneth H., 457
 Leibowitz, Herschel W., 271
 Locke, John L., 63
 LoLordo, Vincent M., 152
 MacKay, Donald G., 31
 Magnusson, David, 28
 Marks, Lawrence E., 407
 Maus, Michael, 217
 Maxwell, O. Dale, 99
 Medin, Douglas L., 117
 Michon, J. A., 304
 Millward, Richard B., 81
 Mueller, John H., 69
 Murch, Gerald M., 241
 Natsoulas, Thomas, 299
 Oldfield, R. C., 133
 Orr, William F., 99
 Parrish, Michael, 271
 Purcell, Dean G., 304, 305
 Reber, Arthur S., 81
 Richardson, Jack, 519
 Robbins, Donald, 303
 Rogers, Michael, 549
 Schneider, Robert J., 549
 Schwartz, Robert M., 189
 Shaughnessy, John J., 499
 Shaver, Phillip, 297
 Smith, Olin W., 307
 Smith, Patricia C., 307
 Sowder, Calvin D., 39
 Spencer, Terry J., 535
 Staddon, J. E. R., 603
 Stadler, Michael, 351
 Stanton, Sara K., 519
 Stevens, Joseph C., 407
 Stilson, Donald W., 549
 Stromeyer, C. F., III, 227
 Sturm, Richard D., 201
 Swartz, Jon D., 307
 Turnure, Cynthia, 339
 Underwood, Benton J., 499
 Vitz, Paul C., 456
 Wallace, William P., 311
 Walsmith, Charles W., 549
 Weir, Morton W., 261
 White, K. G., 421
 Whitman, Charles P., 397
 Wiant, Joseph, 431
 Wilkinson, William C., 237
 Wolf, Gerrit, 153
 Woods, Stephen C., 305, 615
 Zechmeister, Eugene B., 425
 Zimmerman, Joel, 499
 Zimmerman, Robert R., 614

In Memoriam

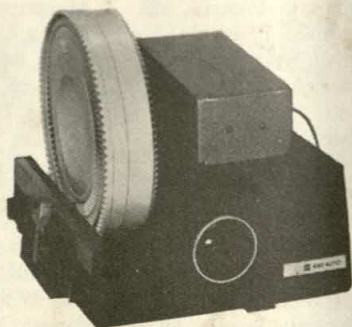
- Karl M. Dallenbach, 463
 Frederic Charles Bartlett, 133
 Gösta Ekman, 295
 Clarence Henry Graham, 291

NEW PHOTO-ELECTRIC PURSUIT ROTOR



The new Pursuit Rotor is manufactured in two models, the MK3 and the MK3/T which incorporate a digital 'time on target'. The variable speed range is from 5 to 50 revolutions per minute and the sampling periods are standardised on 5 — 10 and 20 seconds for simplicity of operation. The MK3/T, as depicted, has the above features and also includes a 4 digit totalising counter to eliminate the use of an external 'total time on target' counter.

SINGLE CHANNEL GROUP TACHISTOSCOPE



The instrument is a general purpose slide projector with remote control facilities and interchangeable lenses. The projector has been modified by the addition of a high-speed galvanometer which operates a shutter membrane to obscure the light path. The rise time of the shutter is 1.5 milli-seconds. This instrument can be used in word, picture and symbol recognition and subliminal research.



Forth Instruments Ltd.,
Engine Road, Loanhead, Midlothian,
Scotland.
Telephone: Loanhead 104.

Manufacturers of Single Channel Group Tachistoscopes, Pulse Generators, Memory Drums, Counter Timers, Voice Switches, Micro Electrode Pullers, Reaction Testers.

NEW

PAINTS

FOR THE

PAINTING OF WALLS AND CEILINGS
AND FOR THE PAINTING OF
WOODWORK AND IRONWORK
AND FOR THE PAINTING OF
METALS AND STONEWORK
AND FOR THE PAINTING OF
FABRICS AND LEATHERWORK
AND FOR THE PAINTING OF
PAPER AND GLASSWORK
AND FOR THE PAINTING OF
CLOTHS AND CURTAINS
AND FOR THE PAINTING OF
SHOES AND HATS
AND FOR THE PAINTING OF
TOYS AND GAMES
AND FOR THE PAINTING OF
FURNITURE AND DECORATION

SINCE

1871

THE

PAINTS

FOR

THE

PAINTING

OF

WALLS

AND

CEILINGS

AND

FOR

